PET-NECK

Mehanna, Hisham; McConkey, Chris C; Rahman, Joy K; Wong, Wai-Lup; Smith, Alison F; Nutting, Chris; Hartley, Andrew Gj; Hall, Peter; Hulme, Claire; Patel, Dharmesh K; Zeidler, Sandra Ventorin von; Robinson, Max; Sanghera, Bal; Fresco, Lydia; Dunn, Janet A

DOI:
10.3310/hta21170

HEALTH TECHNOLOGY ASSESSMENT

VOLUME 21 ISSUE 17 APRIL 2017
ISSN 1366-5278

PET-NECK: a multicentre randomised Phase III non-inferiority trial comparing a positron emission tomography–computerised tomography-guided watch-and-wait policy with planned neck dissection in the management of locally advanced (N2/N3) nodal metastases in patients with squamous cell head and neck cancer

Hisham Mehanna, Chris C McConkey, Joy K Rahman, Wai-Lup Wong, Alison F Smith, Chris Nutting, Andrew GJ Hartley, Peter Hall, Claire Hulme, Dharmesh K Patel, Sandra Ventorin von Zeidler, Max Robinson, Bal Sanghera, Lydia Fresco and Janet A Dunn

NHS
National Institute for Health Research

DOI 10.3310/hta21170

PET-NECK: a multicentre randomised Phase III non-inferiority trial comparing a positron emission tomography–computerised tomography-guided watch-and-wait policy with planned neck dissection in the management of locally advanced (N2/N3) nodal metastases in patients with squamous cell head and neck cancer

Hisham Mehanna,[1]* Chris C McConkey,[2] Joy K Rahman,[2] Wai-Lup Wong,[3] Alison F Smith,[4] Chris Nutting,[5] Andrew GJ Hartley,[6] Peter Hall,[4] Claire Hulme,[4] Dharmesh K Patel,[2] Sandra Ventorin von Zeidler,[7] Max Robinson,[8] Bal Sanghera,[3] Lydia Fresco[9] and Janet A Dunn[2]

[1]Institute of Head & Neck Studies and Education, University of Birmingham, Birmingham, UK
[2]Warwick Clinical Trials Unit, University of Warwick, Coventry, UK
[3]Paul Strickland Scanner Centre, Mount Vernon Hospital, Northwood, UK
[4]Academic Unit of Health Economics, University of Leeds, Leeds, UK
[5]Royal Marsden Hospital, London, UK
[6]University Hospital Birmingham, Birmingham, UK
[7]Department of Pathology, Federal University of Espírito Santo, Vitória, Brazil
[8]Centre for Oral Health Research, Newcastle University, Newcastle upon Tyne, UK
[9]University Hospitals Coventry and Warwickshire, Coventry, UK

*Corresponding author

Declared competing interests of authors: Hisham Mehanna reports that he is a member of the Health Technology Assessment Clinical Trials Board and Claire Hulme reports that she is a member of the Health Technology Assessment Commissioning Board. Andrew GJ Hartley and Max Robinson report grants during the conduct of the study.

Published April 2017
DOI: 10.3310/hta21170

This report should be referenced as follows:

Mehanna H, McConkey CC, Rahman JK, Wong W-L, Smith AF, Nutting C, *et al*. PET-NECK: a multicentre randomised Phase III non-inferiority trial comparing a positron emission tomography–computerised tomography-guided watch-and-wait policy with planned neck dissection in the management of locally advanced (N2/N3) nodal metastases in patients with squamous cell head and neck cancer. *Health Technol Assess* 2017;**21**(17).

Health Technology Assessment is indexed and abstracted in *Index Medicus*/MEDLINE, *Excerpta Medica*/EMBASE, *Science Citation Index Expanded* (SciSearch®) and *Current Contents®*/ Clinical Medicine.

Health Technology Assessment

HTA/HTA TAR

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Impact factor: 4.058

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the ISI Science Citation Index.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: journals.library@nihr.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: http://www.nets.nihr.ac.uk/programmes/hta

This report

The research reported in this issue of the journal was funded by the HTA programme as project number 06/302/129. The contractual start date was in April 2007. The draft report began editorial review in July 2015 and was accepted for publication in June 2016. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

Abstract

PET-NECK: a multicentre randomised Phase III non-inferiority trial comparing a positron emission tomography–computerised tomography-guided watch-and-wait policy with planned neck dissection in the management of locally advanced (N2/N3) nodal metastases in patients with squamous cell head and neck cancer

Hisham Mehanna,[1]* Chris C McConkey,[2] Joy K Rahman,[2] Wai-Lup Wong,[3] Alison F Smith,[4] Chris Nutting,[5] Andrew GJ Hartley,[6] Peter Hall,[4] Claire Hulme,[4] Dharmesh K Patel,[2] Sandra Ventorin von Zeidler,[7] Max Robinson,[8] Bal Sanghera,[3] Lydia Fresco[9] and Janet A Dunn[2]

[1]Institute of Head & Neck Studies and Education, University of Birmingham, Birmingham, UK
[2]Warwick Clinical Trials Unit, University of Warwick, Coventry, UK
[3]Paul Strickland Scanner Centre, Mount Vernon Hospital, Northwood, UK
[4]Academic Unit of Health Economics, University of Leeds, Leeds, UK
[5]Royal Marsden Hospital, London, UK
[6]University Hospital Birmingham, Birmingham, UK
[7]Department of Pathology, Federal University of Espírito Santo, Vitória, Brazil
[8]Centre for Oral Health Research, Newcastle University, Newcastle upon Tyne, UK
[9]University Hospitals Coventry and Warwickshire, Coventry, UK

*Corresponding author H.Mehanna@bham.ac.uk

Background: Planned neck dissection (ND) after radical chemoradiotherapy (CRT) for locally advanced nodal metastases in patients with head and neck squamous cell carcinoma (HNSCC) remains controversial. Thirty per cent of ND specimens show histological evidence of tumour. Consequently, a significant proportion of clinicians still practise planned ND. Fludeoxyglucose positron emission tomography (PET)–computerised tomography (CT) scanning demonstrated high negative predictive values for persistent nodal disease, providing a possible alternative paradigm to ND. Evidence is sparse and drawn mainly from retrospective single-institution studies, illustrating the need for a prospective randomised controlled trial.

Objectives: To determine the efficacy and cost-effectiveness of PET–CT-guided surveillance, compared with planned ND, in a multicentre, prospective, randomised setting.

Design: A pragmatic randomised non-inferiority trial comparing PET–CT-guided watch-and-wait policy with the current planned ND policy in HNSCC patients with locally advanced nodal metastases and treated with radical CRT. Patients were randomised in a 1 : 1 ratio. Primary outcomes were overall survival (OS) and cost-effectiveness [incremental cost per incremental quality-adjusted life-year (QALY)]. Cost-effectiveness was assessed over the trial period using individual patient data, and over a lifetime horizon using a decision-analytic model. Secondary outcomes were recurrence in the neck, complication rates and quality

of life. The recruitment of 560 patients was planned to detect non-inferior OS in the intervention arm with a 90% power and a type I error of 5%, with non-inferiority defined as having a hazard ratio (HR) of no higher than 1.50. An intention-to-treat analysis was performed by Cox's proportional hazards model.

Settings: Thirty-seven head and neck cancer-treating centres (43 NHS hospitals) throughout the UK.

Participants: Patients with locally advanced nodal metastases of oropharynx, hypopharynx, larynx, oral or occult HNSCC receiving CRT and fit for ND were recruited.

Intervention: Patients randomised to planned ND before or after CRT (control), or CRT followed by fludeoxyglucose PET–CT 10–12 weeks post CRT with ND only if PET–CT showed incomplete or equivocal response of nodal disease (intervention). Balanced by centre, planned ND timing, CRT schedule, disease site and the tumour, node, metastasis stage.

Results: In total, 564 patients were recruited (ND arm, $n = 282$; and surveillance arm, $n = 282$; 17% N2a, 61% N2b, 18% N2c and 3% N3). Eighty-four per cent had oropharyngeal cancer. Seventy-five per cent of tested cases were p16 positive. The median time to follow-up was 36 months. The HR for OS was 0.92 [95% confidence interval (CI) 0.65 to 1.32], indicating non-inferiority. The upper limit of the non-inferiority HR margin of 1.50, which was informed by patient advisors to the project, lies at the 99.6 percentile of this estimate ($p = 0.004$). There were no differences in this result by p16 status. There were 54 NDs performed in the surveillance arm, with 22 surgical complications, and 221 NDs in the ND arm, with 85 complications. Quality-of-life scores were slightly better in the surveillance arm. Compared with planned ND, PET–CT surveillance produced an incremental net health benefit of 0.16 QALYs (95% CI 0.03 to 0.28 QALYs) over the trial period and 0.21 QALYs (95% CI –0.41 to 0.85 QALYs) over the modelled lifetime horizon.

Limitations: Pragmatic randomised controlled trial with a 36-month median follow-up.

Conclusions: PET–CT-guided active surveillance showed similar survival outcomes to ND but resulted in considerably fewer NDs, fewer complications and lower costs, supporting its use in routine practice.

Future work: PET–CT surveillance is cost-effective in the short term, and long-term cost-effectiveness could be addressed in future work.

Trial registration: Current Controlled Trials ISRCTN13735240.

Funding: This project was funded by the National Institute for Health Research (NIHR) Health Technology Assessment programme and will be published in full in *Health Technology Assessment*; Vol. 21, No. 17. See the NIHR Journals Library website for further project information.

Contents

List of tables

List of figures

List of boxes

List of abbreviations

AIC	Akaike information criterion
ARSAC	Administration of Radioactive Substances Advisory Committee
BIC	Bayesian information criterion
CI	confidence interval
CR	complete response
CRT	chemoradiotherapy
CT	computerised tomography
DF	disease free
DM	decision model
DR	distant recurrence
eMIT	*Drugs and Pharmaceutical Electronic Market Information*
EORTC	European Organisation for Research and Treatment of Cancer
EQ-5D	EuroQoL-5 Dimensions
EUA	examination under anaesthesia
GP	general practitioner
H&N	head and neck
HNSCC	head and neck squamous cell carcinoma
HPV	human papillomavirus
HR	hazard ratio
ICER	incremental cost-effectiveness ratio
IDSMC	Independent Data and Safety Monitoring Committee
IMRT	intensity-modulated radiotherapy
INB	incremental net health benefit
IQR	interquartile range
LR	local recurrence
MDADI	MD Anderson Dysphagia Inventory
MDM	multidisciplinary meeting
MDT	multidisciplinary team
MRI	magnetic resonance imaging
NB	net health benefit
ND	neck dissection
NICE	National Institute for Health and Care Excellence
N stage	nodal stage
OPC	oropharyngeal cancer
OS	overall survival
PET	positron emission tomography
PSS	personal social services
PSSRU	Personal Social Services Research Unit
QA	quality assurance
QALY	quality-adjusted life-year
QLQ-C30	Quality of Life Questionnaire for Cancer (with 30 questions)
QLQ-H&N35	Quality of Life Questionnaire for Cancer head and neck module with 35 questions
SAE	serious adverse event
TPF	docetaxel, platinum and 5-fluorouracil
TNM	tumour, node, metastases
T stage	tumour stage
UHB	University Hospital Birmingham
UHCW	University Hospitals Coventry and Warwickshire
VAS	visual analogue scale
WT	within trial

Plain English summary

What was the problem?

Head and neck cancer has devastating effects on patients' self-image, speech and swallowing. Chemoradiotherapy (CRT) has become an important way to treat this cancer. In patients whose cancer has spread to the neck lymph glands, current treatment includes removal of the neck lymph glands using an operation called neck dissection (ND). This can have significant complications and after-effects, such as shoulder disability, disfigurement of the mouth and neck, and long-term pain.

With the improvement in CRT, some now believe that ND may no longer be required if the neck disease is treated adequately by CRT. Furthermore, owing to an improved scanning technology called positron emission tomography (PET)-computerised tomography (CT), there is now better ability to identify patients whose neck disease has responded completely to CRT and who do not require a ND.

What did we do?

We compared routine ND with a PET–CT-guided watch-and-wait policy in patients with advanced neck disease to ascertain if PET–CT would result in a survival rate similar to ND, while reducing the number of NDs being performed. We also looked at the costs of both treatment strategies and their impacts on patients' quality of life.

What did we find?

Patients who received PET–CT-guided surveillance showed similar survival outcomes to those who received planned ND. PET–CT surveillance also resulted in fewer complications and lower costs, supporting its use in routine practice.

Scientific summary

Background

Chemoradiotherapy (CRT) has become the preferred method of treatment for patients with advanced head and neck squamous cell carcinoma (HNSCC). The traditional standard care in the UK for these patients included undertaking a neck dissection (ND) (surgery to remove the lymph nodes in the neck) before or after CRT. However, there is considerable debate about whether or not ND is actually needed or whether or not CRT alone is sufficient to treat the disease without the need for surgery and its added complications. The standard imaging technology for assessing response to CRT has been computerised tomography (CT) and/or magnetic resonance imaging (MRI). However, more advanced functional modalities [especially positron emission tomography (PET) scans] have in recent decades been shown to have a high negative predictive value for assessing response. Using a combination of PET with CT, for example, has been shown in retrospective studies to have a higher predictive value than CT or MRI combined, making it possible to perform a ND only if the nodal response to treatment is incomplete. There is, however, a lack of multicentre high-quality evidence.

Objectives

- To compare the efficacy of a PET–CT-guided active surveillance (watch-and-wait) policy with the current practice of planned ND on overall survival (OS), disease-specific survival, recurrence, quality of life and cost-effectiveness in the management of advanced (N2 or N3) nodal metastasis in patients treated with CRT for their HNSCC primary.
- To assess the predictive value of PET–CT scanning in detecting persistent/residual disease in the primary site of patients with HNSCC treated with primary CRT.

End points

Primary end point

- Overall survival at 2 years.
- Cost-effectiveness [incremental cost per incremental quality-adjusted life-year (QALY)].

Secondary end points

- Disease-specific survival.
- Recurrence in the neck.
- Quality of life.
- Complication rates.
- Accuracy of PET–CT scanning for assessing the primary tumour.

Study design and methodology

A two-arm pragmatic multicentre randomised non-inferiority trial was performed to compare a PET–CT-guided watch-and-wait policy (experimental arm) with the current planned ND policy (control arm) in HNSCC patients with advanced neck metastasis treated by radical CRT. A total of 564 patients were randomised in a 1 : 1 ratio.

Stratification was performed according to centre, timing of ND (before vs. after CRT), chemotherapy schedule [concomitant platinum, concomitant cetuximab Erbitax® (Merck Biopharma, Darmstadt, Germany), neoadjuvant platinum, neoadjuvant docetaxel Taxotere® (Sanoti-Aventis, Gentilly, France), platinum and 5-fluorouracil (TPF)], disease site (oropharyngeal, laryngeal, oral, hypopharyngeal or occult), tumour (T) stage (T1–T2 vs. T3–T4 vs. occult) and nodal (N) stage (N2a–N2b vs. N2c–N3).

Treatment and investigations, radiotherapy and chemotherapy protocols

For each patient, the participating centre decided on the CRT schedule, which was chosen from an approved list of schedules. All approved schedules were standard normal schedules used in the UK. All were supported by a strong evidence base, and all were considered biologically equivalent.

Post-chemoradiotherapy assessment

This was performed at 12 (9–13) weeks after completion of CRT.

Patients were assessed for response to the CRT by:

- control arm – a single CT/MRI scan and examination (clinical or under anaesthetic)
- experimental arm – a single PET–CT scan followed by examination (clinical or under anaesthetic).

Diagnostic criteria and reporting protocols for PET–CT scanning

Standardised criteria for reporting of PET–CT scans were disseminated to all participating centres. A core laboratory facility was set up in the Paul Strickland Scanner Centre, Mount Vernon Hospital, to read scans for units that had the equipment and ability to perform PET–CT but did not have the expertise to read them. The laboratory also performed second-stage quality assurance on all PET–CT scans performed for study patients.

Type of neck dissection

Modified radical ND involving nodal levels I to V or selective NDs were acceptable provided that involved nodal groups were included.

Timing of neck dissection

Neck dissection before CRT had to be performed within 4 weeks of randomisation. ND after CRT had to be performed 4–8 weeks after completion of CRT.

Sample size determination

The study was planned to randomise 560 patients (280 to PET–CT surveillance and 280 to planned ND), which would allow for the demonstration of non-inferiority of the PET–CT arm, with a 5% one-sided significance and 90% power, defining non-inferiority as no worse than 10% below the estimated 75% 2-year OS of the control arm, that is, having a hazard ratio (HR) no higher than 1.50. This allowed for a 3% loss to follow-up.

Follow-up

Follow-up was at 6, 12 and 24 months post randomisation and continued until at least 24 months after randomisation. Long-term health status data on death and recurrence were collected for patients until the end of the study. Patients were flagged with the Office for National Statistics and copies of their death certificates were requested for long-term follow-up. This will be reported in a long-term follow-up paper.

Key inclusion/exclusion criteria

Inclusion criteria

Patients with *all* of the criteria listed below were eligible:

- histological diagnosis of oropharyngeal, laryngeal, oral, hypopharyngeal or occult HNSCC
- clinical and CT/MRI imaging evidence of nodal metastases staged N2 (a, b or c) or N3

- indication to receive curative radical concurrent CRT for primary
- fitness for ND surgery
- ND was technically feasible to perform and to remove nodal disease (e.g. no carotid encasement, no direct extension between tumour and nodal disease)
- aged 18 years old or more
- able to give informed consent
- receiving one of the CRT regimens approved by the study.

Exclusion criteria

Patients with *any* of the criteria listed below were ineligible:

- undergoing resection for their primary tumour, for example resection of the tonsil or base of tongue with flap reconstruction (diagnostic tonsillectomy was not considered an exclusion criteria)
- distant metastases to chest, liver, bones or other sites
- previous treatment for HNSCC
- pregnant
- had had another cancer diagnosis in the past 5 years (except basal cell carcinoma or carcinoma of the cervix in situ).

Patients with N2 or N3 histologically and/or cytologically proven squamous cell carcinoma and an occult primary (after examination under anaesthetic and PET–CT scan) were eligible for the PET-NECK trial if they were going to be treated with CRT.

Patients undergoing neoadjuvant chemotherapy followed by concomitant CRT were eligible for the PET-NECK trial. If these patients were randomised to the ND (control) arm, it was recommended that they have a ND after, not before, CRT. Patients with recurrence remained in the trial for the purposes of follow-up and data collection.

Results

In total, 564 patients were recruited (ND arm, $n = 282$ and surveillance arm, $n = 282$; 17% N2a, 61% N2b, 18% N2c and 3% N3). Eighty-four per cent had oropharyngeal cancer. Seventy-five per cent of tested cases were p16 positive. The median length of follow-up was 36 months.

The HR for OS was 0.92 [95% confidence interval (CI) 0.65 to 1.32] indicating non-inferiority. The upper limit of the non-inferiority HR margin of 1.50, which was informed by patient advisors to the project, lies at the 99.6 percentile of this estimate ($p = 0.004$). There were no differences in this result by p16 status. There were 54 NDs performed in the surveillance arm, with 22 surgical complications, and 221 NDs in the ND arm, with 85 complications. Quality-of-life scores were slightly better in the surveillance arm. Compared with planned ND, PET–CT surveillance produced an incremental net health benefit of 0.16 QALYs (95% CI 0.03 to 0.28 QALYs) over the trial period, and 0.21 QALYs (95% CI to 0.41 to 0.85 QALYs) over the modelled lifetime horizon.

Conclusions

Positron emission tomography–computerised tomography-guided active surveillance showed similar survival outcomes to the ND arm, but resulted in considerably fewer NDs, fewer complications and, probably, lower costs. Further exploration of the significance of persistent nodal enlargement but no PET uptake is required.

Trial registration

This trial is registered as ISRCTN13735240.

Funding

Funding for this study was provided by the Health Technology programme of the National Institute for Health Research.

Chapter 1 Introduction

Background

Head and neck squamous cell carcinoma (HNSCC) is the sixth most common cancer worldwide, with approximately 500,000 new cases per year.[1] It poses a significant therapeutic problem and has a high mortality and morbidity. Survival rates [apart from those for oropharyngeal cancer (OPC)] have not considerably improved over the past three decades despite newer aggressive surgical and chemoradiotherapy (CRT) regimens. Furthermore, both the disease and its treatments have considerable effects on vital functions (such as speech, eating, swallowing and appearance) and can result in significant functional deficits and quality-of-life effects.

Organ preservation treatment protocols, using radical concomitant CRT, have evolved in the past three decades, resulting in improved control rates (6.5% absolute improvement) at the primary site compared with radiotherapy alone.[2–4] Therefore, in many centres, CRT has become the preferred first line of treatment for several types of HNSCC, especially of the base of the tongue, tonsil, larynx and hypopharynx.

For patients who have advanced nodal metastasis in the neck [tumour, node, metastases (TNM) stages N2 or N3 nodes > 3 cm in size], the evidence for management is controversial.[1] Previously, the standard care in these patients was to perform a neck dissection (ND) (an operation to remove all lymph glands in the neck), either before or after CRT. A ND can result in morbidity, which can be lifelong, and (even a small risk of) mortality.[4] Since 2000, it appears that there has been a shift towards active surveillance guided by imaging, but there remains a significant proportion of patients being treated by planned routine ND.[5] Therefore, there continues to be lack of consensus regarding the best management for advanced nodal disease in patients receiving CRT.

This controversy continues mainly because of the poor quality and contradictory evidence (level 3/4) from prospective and retrospective case series for both management strategies. Furthermore, the advent of newer and more accurate functional modalities for the detection of persistent disease [such as positron emission tomography (PET)–computerised tomography (CT) scanning[6,7]] has further strengthened this debate. Several studies and systematic reviews have reported a high negative predictive value for PET scanning for the detection of persistent nodal disease. However, studies are small, mainly retrospective, single-centre studies. Many authorities on head and neck (H&N) cancer and literature reviews[7,8] have stressed the need for a multicentre randomised trial to obtain an answer to this important question.

Existing research

Evidence in support of planned (routine) neck dissection

Until recently the standard care in the UK was to perform a ND with CRT. In other countries, a significant proportion of patients are still treated with planned ND.[5] Proponents of this management policy maintain that CRT does not eradicate large-volume nodal disease in a large proportion of patients (up to 50%), putting them at risk of recurrence. Some believe that for most of these patients, salvage by surgery will not be possible,[9] resulting in devastating consequences.

The only randomised controlled trial available in the literature on the subject compared conservative follow-up with planned ND before radical CRT.[10] It found that there was significantly improved disease-specific survival following ND. However, it was small (50 patients) and had several major limitations, casting strong doubts on the validity of its results and findings. Other single-arm and retrospective studies have reported that planned

ND demonstrated excellent locoregional control, although some did not detect improvement in overall or disease-specific survival in their series.[11–14]

In addition, it has been found that in up to 40% of patients who show a complete clinical response to CRT but who undergo ND 8–10 weeks later tumour deposits can still be detected histologically in the ND specimen.[11,14,15] It has therefore been suggested that CRT does not completely eradicate the tumour in up to 40% of patients. Furthermore, proponents of ND maintain that selecting patients at a high risk of persistent disease is not possible using CT and/or magnetic resonance imaging (MRI), as studies have found that clinical and imaging evidence of complete response (CR) of nodal disease to CRT does not predict a complete pathological response (i.e. it does not correlate with an absence of pathological evidence of the disease in the ND specimen).[12,14]

Evidence supporting a watch-and-wait policy in patients with complete response to chemoradiotherapy

Many clinicians now advocate a conservative surveillance policy, performing ND only if there is clinical evidence of persistent nodal disease after CRT. The rationale is that, among patients who exhibit a clinical CR in the neck following CRT, the recurrence rate in the neck is low (< 10%), and similar to that (8%) in patients with pathologically negative neck following ND.[8,16] Other level 2/3 studies[9,17,18] have also found that ND in patients who show a CR to CRT does not confer any benefit in terms of improved survival or reduction in nodal recurrence compared with a watch-and-wait policy. Importantly, a large retrospective cohort study found that 43% of patients showed CR on CT, with a control rate of 92% at 5 years. Among those who experienced less than a CR, the 5-year control rate among those who underwent ND was similar, at 90%, but among those who did not undergo ND it was significantly lower (76%).[19]

There is also evidence to suggest that, in many cases, the cells in the residual nodal deposits found on histology in ND specimens following CRT are not viable.[7,18] An experimental study, examining a proliferation marker, Ki-67, appears to confirm this hypothesis.[20]

Accuracy of PET–CT scanning in the assessment of response to chemoradiotherapy and detection of residual nodal disease

Crucial to a watch-and-wait policy is the ability to detect residual disease in the neck post CRT, so that these patients can be targeted to have a ND. Detection of residual disease in the neck is not accurately assessed by examination in the clinic or by CT and/or MRI. Evidence suggests that PET is able to accurately identify those patients who do not have residual tumour in their neck following CRT (i.e. PET has a high negative predictive value).

Positron emission tomography has been reported in several small single-institution prospective and retrospective series and two meta-analyses[21,22] to have a negative predictive value of 90–100% for the detection of persistent nodal metastasis following radiotherapy or CRT, with a variable positive predictive value of approximately 30–60%.[23–31] It has also been shown in several studies to have higher predictive values than CT and/or MRI,[21,27,32,33] and better than combined clinical examination and ultrasound.[31] Finally, studies have found that co-registering PET and CT (PET–CT) is even more accurate than PET alone[33,34] (*Table 1*).

TABLE 1 [a]Reported accuracy of PET–CT, PET and CT[33]

Modality	Sensitivity (%)	Specificity (%)	Accuracy (%)	Positive predictive value (%)	Negative predictive value (%)
PET–CT	98	92	94	88	99
PET	87	91	90	85	92
CT	74	75	74	63	83

a $n = 125$ lesions in 64 patients.

There is also evidence from same studies that PET–CT is highly sensitive to the detection of residual tumour at the primary site following CRT and radiotherapy, with a high negative predictive value of 90–100%.[35–37] Indeed, it has been suggested that PET–CT may have similar a predictive value in the exclusion of residual disease at the primary site to examination under anaesthesia (EUA), the current gold standard, and may be capable of replacing it as a result of it being less invasive.[38,39] However, there is no existing level 1 evidence to corroborate this theory.

The study had not intended to measure human papillomavirus (HPV) status at the time of inception. However, over the past 5 years, there have been reports on the significant increasing incidence of HPV-associated OPC, with the proportion of OPCs that are HPV associated increasing from 40.5% to 72% over a period of 20 years.[40] HPV association appears to have a remarkable effect on the prognosis and outcomes of treatment, with a 58% reduction in the risk of death [hazard ratio (HR) 0.42, 95% confidence interval (CI) 0.27 to 0.66].[41] This has necessitated examining outcomes by HPV status.[42]

Research objectives

- To test the hypotheses that a PET–CT-guided watch-and-wait policy (experimental arm) is non-inferior to the current practice of planned ND (control arm) when comparing overall survival (OS) in the management of advanced (N2 or N3) nodal metastasis in patients treated with CRT for primary HNSCC.
- To assess the cost-effectiveness of the PET–CT-guided watch-and-wait policy.
- To compare, as secondary outcomes, the efficacy of the PET–CT-guided watch-and-wait policy in terms of disease-specific survival, recurrence and quality of life.
- To describe the accuracy of PET–CT in the detection of persistent disease at the primary site and the neck following CRT.

Chapter 2 Methods

Trial design

The trial was a two-arm, pragmatic, multicentre, randomised, non-inferiority trial comparing a PET–CT-guided watch-and-wait policy (experimental arm) with the current planned ND policy (control arm) in HNSCC patients with advanced neck metastases and treated with radical CRT.

Target recruitment was 560 patients. Patients were randomised in a 1 : 1 ratio. Primary outcomes were OS and health economics. Secondary outcomes were recurrence in the neck, complication rates and quality of life.

Amendments to the protocol

Substantial amendments to the trial protocol were approved since conception, which were significant to evaluate the economic costs, to help increase recruitment and to clarify eligibility and which included the following:

- The introduction of health economics evaluation at local sites. A health economic cost analysis was required because the avoidance of a ND with its costs and possible complications and morbidity is expected to be one of the most significant benefits of the watch-and-wait policy. Furthermore, the economic effect of replacing EUA and CT with a PET–CT scan would need to be evaluated to assess its potential for the future. The aim of the economic evaluation was to identify the within-trial (WT) and long-term incremental cost-effectiveness of PET–CT-guided watch-and-wait compared with planned ND in HNSCC patients.
- To extend the inclusion criteria to allow patients with occult primary tumours to enter the PET-NECK trial. Patients occasionally have pathologically occult tumours but, if all other diagnostic procedures point towards a H&N primary, and all other primary sites are excluded, they are routinely diagnosed and treated as having H&N cancer. There was no reason why these patients should not also have been given the opportunity to enter the PET-NECK trial if they wished and if they met all other criteria.
- To allow ND to be performed before or after CRT in patients randomised to receive standard ND (control arm). There had been a change in practice in the previous 2–3 years regarding the timing of planned ND. In the USA and Europe, planned ND is usually carried out after CRT, rather than before, and there has been a gradual shift towards this practice in the UK over the previous 2–3 years. Therefore, amending the control arm to allow this was intended to increase recruitment and 'future-proof' the study if this trend continues in the UK.
- To exclude patients with tumours that were not histologically squamous cell carcinoma and patients with primary nasopharyngeal carcinoma. Both non-squamous cell carcinomas and nasopharyngeal carcinomas have different biological behaviours and natural histories from squamous cell carcinoma of the H&N, and they should not be treated in a similar manner. Nasopharyngeal carcinoma is highly sensitive to radiotherapy and should not be treated by a ND. Non-squamous cell carcinoma is much less radiosensitive and, therefore, a ND is indicated in these cases. The literature available pertains to only squamous cell carcinoma in sites of the H&N other than nasopharyngeal.
- To clarify that patients with equivocal PET–CT scans should receive a ND. The management policy in the UK for patients with equivocal PET–CT scans at the time of conducting the study was that they should receive a ND.

Ethics and research and development approvals

A favourable opinion was given by the Oxfordshire Multi-Research Ethics Committee in May 2007 (reference number 07/Q1604/35). Research and development approval was obtained from University Hospitals Coventry and Warwickshire (UHCW) in June 2007, and permissions to conduct the study at each

site were obtained from the NHS trusts covering the Heart of England, Coventry and Warwickshire, Bath, Velindre, London North West, London Central, London South East, Marsden, Poole, Barnet and Chase Farm, Blackpool, Lancashire, Beatson West of Scotland, Abertawe Bro Morgannwg, Sunderland, Luton and Dunstable, Hull and East Yorkshire, Basildon and Thurrock, Southend, Sheffield and Chesterfield, Bradford, Nottingham, East Lancashire, Mid Essex, Aintree, Derby, Royal Devon, Bristol North, Bristol South, East Kent, Portsmouth, Plymouth, Manchester Central, Christie, Pennine, Hertfordshire, Dumfries and Galloway, Lothian, Royal Surrey, Newcastle Upon Tyne, Dudley, Grampian, Wolverhampton, Belfast, South Devon, Gloucestershire, Walsall, Birmingham West Midlands and South Tees. Participating hospitals are listed in the *Acknowledgements* of this report. All sites were activated between September 2007 and May 2012.

Sponsorship

The PET-NECK trial was co-ordinated by the Warwick Clinical Trials Unit and it has a co-sponsorship agreement with UHCW. The Warwick Clinical Trials Unit stores the data within the University of Warwick central information technology service host computers. All data are securely stored under the Data Protection Act 2004[43] and adhere to the University of Warwick Clinical Trials Unit data sharing standard operating procedure in which data sharing agreements have to be approved by both the Trial Management Group and the sponsor.

Participants

The study sought to recruit patients diagnosed with HNSCC with advanced nodal metastases who had not received previous treatment for their HNSCC and who had had no other cancer diagnoses within the past 5 years (exceptions given in *Exclusion criteria*), from H&N cancer-treating centres throughout UK NHS hospital trusts.

Inclusion criteria

Patients who met *all* of the criteria listed below were eligible to participate in the study:

- had a histological diagnosis of oropharyngeal, laryngeal, oral, hypopharyngeal or occult HNSCC
- had a clinical and CT/MRI imaging evidence of nodal metastases, stage N2 (a, b or c) or N3
- had a multidisciplinary team (MDT) decision to receive curative radical concurrent CRT for primary
- had an indication to receive one of the CRT regimens approved by the study
- were fit for ND surgery
- ND was technically feasible to perform to remove nodal disease (e.g. no carotid encasement, no direct extension between tumour and nodal disease)
- were aged $\geq$ 18 years
- were able to provide written informed consent.

Patients with N2 or N3 histologically and/or cytologically proven squamous cell carcinoma and an occult primary (after EUA and PET–CT scan) were eligible for the trial if they were going to be treated with CRT.

Exclusion criteria

Patients who met *any* of the criteria listed below were ineligible:

- had tumours that were not squamous cell carcinomas histologically
- were undergoing resection for their primary tumour, for example resection of the tonsil or base of tongue with flap reconstruction (diagnostic tonsillectomy was not considered an exclusion criteria)
- had N1 stage nodal metastasis
- were receiving neoadjuvant CRT with no concomitant chemotherapy
- were receiving adjuvant chemotherapy
- were undergoing chemotherapy with or without radiotherapy for palliative purposes

- were undergoing radiotherapy alone (this is not an optimal treatment for neck node disease)
- had distant metastases to the chest, liver, bones or other sites
- were unfit for surgery or CRT
- had received previous treatment for HNSCC
- had primary nasopharyngeal carcinoma
- were pregnant
- had had another cancer diagnosis in the past 5 years (with the exception of basal cell carcinoma or carcinoma of the cervix in situ).

Patients undergoing neoadjuvant chemotherapy followed by concomitant CRT were eligible to enter the trial.

The patients in the study were similar to those in the studies that established the efficacy of CRT of the larynx[42] and oropharynx,[44] and were similar to the patients included in a wide range of studies examining PET–CT in the assessment of persistent nodal disease.[21,22]

Settings and locations

In total, 38 centres (57 NHS hospitals) throughout the UK took part in the study. Participating centres were Department of Health-approved MDTs working in NHS clinics and undertaking the management of H&N cancer. MDTs are assessed according to defined criteria set by the Department of Health, including national standards for diagnosis, staging, therapy, radiology, pathology and patient support.[45] All clinicians have to be core members of the approved MDTs meeting minimum qualifications and throughput criteria. All trusts undertaking the management of H&N cancer undergo regular peer review and quality assurance (QA).[46] To participate, centres had to have centrally satisfactorily conducted a peer review report in the last 2 years. A list of participating hospitals is given in the *Acknowledgements* of this report.

All centres were required to provide confirmation of trust research and development and Administration of Radioactive Substances Advisory Committee (ARSAC) approval to conduct the study at each site. Centres also needed access to a PET–CT scanner either locally or at a distant site.

Administration of Radioactive Substances Advisory Committee approval

Administration of Radioactive Substances Advisory Committee licences were study specific and related only to the PET–CT centre, which was stated on the application.

Therefore, when a licence was held for the PET-NECK trial at a particular PET–CT centre, the site could send their patients to be scanned there as long as the ARSAC licence holder agreed to oversee the safety of the patients. The trials office had a list of ARSAC licence holders and arranged this for the centre. However, when PET–CT centres were not close by or when the site had its own PET–CT centre but did not hold a licence, its radiologist had to make an application under the *Ionising Radiation (Medical Exposure) 2000 Regulations*[47] to the appropriate ARSAC. Approval of applications took around 1 month and patients could not be scanned without the licence being in place. ARSAC licences were issued for 2 years and had to be renewed after that time. The trials office kept a log of all expiry dates for ARSAC certificates and reminded the radiologist to renew them and send a copy of the renewal. A list of ARSAC licence holders is given in the *Acknowledgements* section of this report.

Quality assurance testing for PET–CT scanners

Before sites could scan patients in the surveillance arm of the trial, the PET–CT scanner to be used had to receive a successful QA review by the study's central physicist.

All scanners had to be quality assured by the physicist. Each site had to arrange with the PET–CT centre to send a test case to the PET-NECK trial core laboratory for standardisation and quality control for each scanner that would be used in the trial.

The test case had to be an ^{18}F or ^{68}Ge Dicom phantom scan. All Dicom images had to be anonymised with patient trial number and initials and the compact disc or digital versatile disc had to be correctly labelled with 'PET-NECK Trial', scan date, site name and coded identifier and be sent along with an anonymised copy of the local report and a completed study Dicom patient scan transfer form.

Providers of PET–CT scanning units used in the study were InHealth (InHealth Group, High Wycombe, UK), Alliance (Alliance Healthcare, Hinckley, UK) and Cobalt Healthcare (Cobalt Healthcare, Cheltenham, UK). Scanners used for the study were either 'static', that is, fixed at a specific hospital site, or 'mobile', that is, able to travel various hospitals around the country.

Static scanners had to be quality assured every 2 years; however, mobile scanning units had to be tested by the core laboratory half yearly. The trial team kept a log of the pass dates for scanning units and organised with the PET–CT providers to renew the QA review.

A set-up visit to each centre by the trial co-ordinator also took place, during which information about the trial protocol was presented and site staff were given the opportunity to clarify information. Each site was provided with the TNM staging manual and written procedural requirements for PET–CT scanning and reporting. Once all approvals were in place, an activation letter was sent to the principal investigator and trust research and development department for each site to activate the start of recruitment.

Recruitment procedure

Participants were identified following pre-study assessments for diagnosis and tumour staging. Clinical, radiological and pathological staging was performed in accordance with the Union for International Cancer Control's *TNM Classification of Malignant Tumours* staging manual.[48]

Assessments for diagnosis and tumour staging

Each of the following assessments had to be done within the 4 weeks prior to randomisation, or at least CT/MRI or EUA had to have been carried out within the 4 weeks before randomisation:

1. CT/MRI of the primary site and neck (which had to be carried out before biopsy or not less than 2 weeks after biopsy).
2. Examination of the primary site with biopsy. In most cases, this necessitated examination under general anaesthesia. However, it was not mandatory if adequate clinical examination and biopsy could be performed without general anaesthetic.
3. CT of the chest.
4. Assessment for fitness for anaesthetic by treating clinician or anaesthetist.

Patient treatment was discussed by the MDT and each patient was considered for the PET-NECK trial. Eligible patients were invited to the clinic to discuss treatment and were told about the PET-NECK study. The study was discussed in detail with the patient and carer and all questions were answered. Once a patient decided to enter the study, informed consent was taken.

Informed consent

Adequate time was allowed for patients to consider their participation in the trial. There was no pre-agreed specified time by which to consent. Consent was required to be informed and voluntary, with time for questions and reflection. However, the patient also had the right to make an immediate decision to consent.

Consent to participate in the PET-NECK study was sought by the clinician involved in the patient's care, with the involvement of the research nurse in the consent discussion. Patients were asked to confirm their consent to participate in the PET-NECK trial by initialling the appropriate boxes on the consent form and signing the form in the presence of the person taking consent. A copy was given to the patient, another copy was kept in the patient notes and the original was kept in the local site file.

The PET-NECK trial intervention

Patient randomisation

In total, 564 patients were recruited from 37 centres (43 NHS hospitals). For all patients recruited to the study, written informed consent was obtained.

Randomisation occurred centrally through the Warwick Clinical Trials Unit, where the study was being managed. Treatment allocation was performed using a minimisation algorithm and was stratified by the following: centre, timing of planned ND (before or after CRT), chemotherapy schedule {concomitant platinum, concomitant cetuximab [Erbitux® (Merck Biopharma, Darmstadt, Germany)], neoadjuvant and concomitant platinum, neoadjuvant docetaxel [Taxotere® (Sanofi-Aventis, Gentilly, France)] and platinum and 5-fluorouracil (TPF) with concomitant platinum} disease site (oropharyngeal, laryngeal, oral, hypopharyngeal or occult), tumour (T) stage (T1–T2 vs. T3–T4) and nodal (N) stage (N2a–N2b vs. N2c–N3) (*Figure 1*).

At randomisation, the eligibility of the patient was checked and the trial number and treatment arm allocated. Before being informed of the random allocation, each patient was asked to complete a quality-of-life questionnaire. Following randomisation, the patient's general practitioner (GP) was notified of study participation by letter.

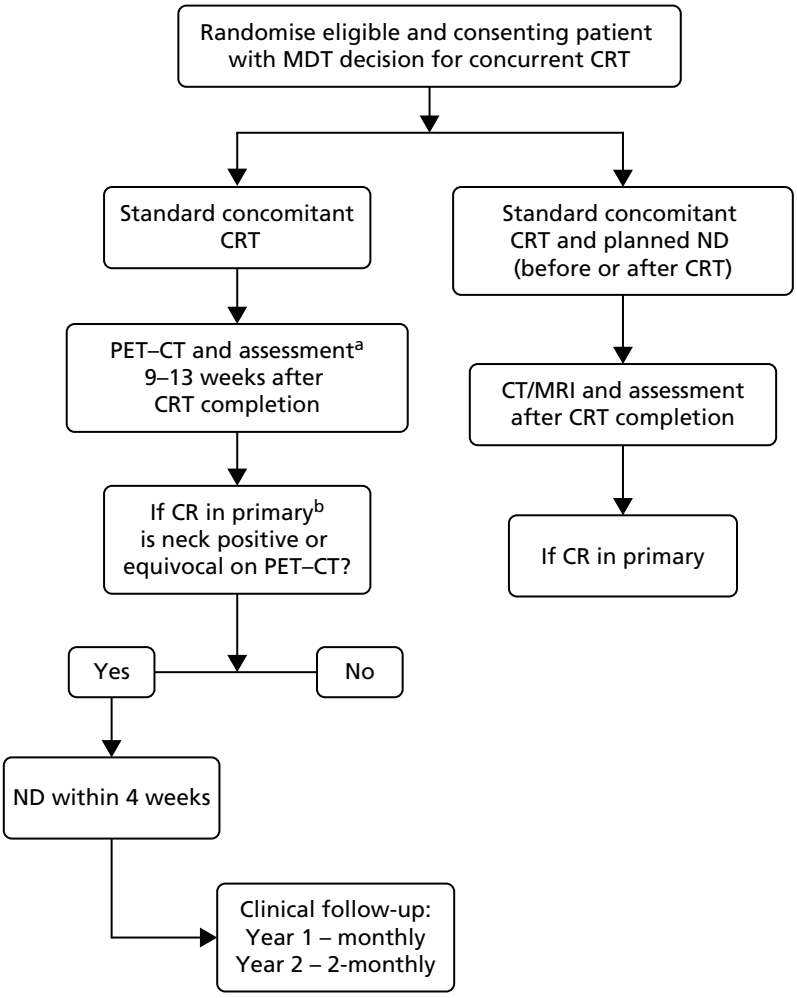


FIGURE 1 Trial schema. a, EUA; b, if the patient's primary tumour showed no response or an incomplete response to CRT or the patient developed a recurrence in primary site or neck, the patient was immediately referred back to the MDT for consideration for salvage surgery.

Treatment

Chemoradiotherapy regimens

To be enrolled in the study, patients had to have been recommended for concomitant CRT by the MDT. Patients randomised to the ND arm underwent ND either within 4 weeks before or within 4–8 weeks after CRT. For each patient, the participating centre had to specify at randomisation the schedule to be used. The CRT schedule was required to be one that the centre uses in its normal peer-reviewed practice *and* on the list of approved trial schedules in *Box 1*.

The recommended CRT schedule and the approved variations to the schedule were selected after consultation with several oncologists nationwide. The approved schedules were selected because they each fulfilled the following criteria: the schedule was supported by a strong evidence base and the schedule was well established and a standard schedule in UK centres. Variations from the approved trial schedules as a result of emerging evidence or changes in the practice of a centre were individually appraised and approved when appropriate by the PET-NECK trial management team. Patients receiving radiotherapy only (even if receiving accelerated radiotherapy schedules) were not eligible to enter the trial.

Planned neck dissections

The randomising centre was required to decide before randomisation whether, in the case of a patient being randomised to the ND (control) arm, the planned ND would be performed before or after CRT.

Timing of neck dissection

If the ND was to be performed before CRT, it had to be performed within 4 weeks of randomisation.

If the ND was to be performed after CRT, it had to be performed 4–8 weeks after completion of CRT. It was recommended that patients undergo assessment by CT when possible to assess the primary site (but this was not mandatory) and clinical examination with or without EUA before ND.

Type of neck dissection

The recommended surgical procedure was a modified radical ND. This involves the removal of lymphatic structures in levels I–V, with preservation of one or more of the following: spinal accessory nerve, internal jugular vein and sternocleidomastoid muscle.[54]

Selective NDs (including lymphatic tissues in fewer than the five levels I–V) were also acceptable provided the following conditions were met:

1. All clinically evident involved nodal groups were included in the ND.
2. The lymphatic tissues of levels I–V that had *not* been removed in the ND were included in the radiotherapy fields to a minimum of 50 Gy.

The type of access incision was left to the surgeon to decide.

Quality assurance of histological assessment of neck dissection specimens

Histological assessment of ND specimens was performed in accordance with the requirements and standards of the minimum pathology data set published by the Royal College of Pathologists.[46]

BOX 1 Approved CRT schedules

Recommended trial chemoradiotherapy regimen

The recommended standard schedule for the trial was:

- radiotherapy doses of 65–70 Gy in 30–35 daily fractions of 2 Gy or more with at least two doses of concomitant 3-weekly i.v. cisplatin (100 mg/m^2) or carboplatin (4.5–5 AUC).[2,3]

The following radiotherapy schedules are examples of approved schedules and are considered equivalent: 70 Gy in 35 fractions and 65 Gy in 30 fractions.

Approved variations to the recommended trial chemoradiotherapy schedule

The following variations to the recommended trial standard regimen were also permitted:

1. Radiotherapy schedule variations: doses of 55 Gy in 20 daily fractions or equivalent.
2. Concomitant chemotherapy schedule variations:
    i. Weekly cisplatin doses (30–40 mg/m^2) to a minimum cumulative cisplatin dose of 200 mg/m^2 or more or weekly carboplatin (1.5 AUC).[49,50]
    ii. The use of epidermal growth factor receptor modulation with cetuximab instead of cisplatin-based chemotherapy was permitted: loading dose of 400 mg/m^2 body surface area followed by weekly i.v. infusions of 250 mg/m^2.[51,52]
    iii. Variations from the approved trial schedules above as a result of emerging evidence or changes in the practice of a centre were at times permitted, after consultation with the PET-NECK trial management team.

Neoadjuvant with concomitant chemotherapy schedules

1. Patients undergoing neoadjuvant chemotherapy followed by concomitant CRT were eligible for the trial. If these patients were randomised to the ND (control) arm, it was recommended that patients have their ND within 2 weeks of randomisation if possible, followed by the neoadjuvant schedule.
2. Neoadjuvant with concomitant chemotherapy schedules could be used at the centre's discretion, provided that the following schedules were used:
    i. Up to three cycles of neoadjuvant platinum-based chemotherapy were permissible, for example cisplatin 75–100 mg/m^2 or carboplatin AUC 5–6 and 5-fluorouracil 1000 mg/m^2 a day for 4–5 days or equivalent.
    ii. Docetaxel 75 mg/m^2 with platinum (75–100 mg/m^2) and 5-fluorouracil (1000 mg/m^2) schedules were also permitted.[53]
3. Variations from the approved trial schedules above, as a result of emerging evidence or changes in the practice of a centre may have been permitted, after consultation with the PET-NECK trial management team.
4. Adjuvant CRT and neoadjuvant-only chemotherapy followed by radiotherapy schedules were not permissible.

AUC, area under the curve; i.v., intravenous.

A central pathology review of 10% of pathology specimens was performed for the purposes of ensuring quality control. The cases were selected at random by the PET-NECK trial office and requested from the participating centres. There were no differences in the results of the local centre pathologist and the central pathology reviewer.

Three-month post-chemoradiotherapy assessment

All patients were assessed for response after CRT. In the case of patients randomised to the PET–CT (surveillance) arm, this assessment took place 9–13 weeks after completion of CRT. In patients who were randomised to ND (control) arm and who received ND before CRT, assessment was performed 9–13 weeks after CRT. In patients in the control arm who received ND after CRT, assessment for response to CRT was done 4 weeks prior to the ND.

The timing of the assessment was chosen to reflect literature findings that PET–CT is most accurate when done at least 8 weeks post CRT, and preferably > 12 weeks after CRT. However, this was balanced by the need to perform the ND as soon as possible to avoid regrowth of the residual tumour post CRT, and the increased technical difficulty in performing ND later than 12 weeks post CRT because of ensuing fibrosis and scarring. As this was a pragmatic trial, some flexibility was afforded.

At any time, if a patient was found to have residual primary disease, he or she was immediately referred back to the treating clinician and MDT for consideration for salvage treatment.

Neck dissections for persistent disease identified on PET–CT post chemoradiotherapy

Persistent disease on PET–CT was defined as metastatic nodal disease in the neck identified by PET–CT at the 3-month post-CRT assessment. Patients in whom PET–CT findings 3 months post CRT were equivocal had to be treated as if they had persistent disease and had to undergo ND. The ND had to be performed within 4 weeks of a decision at the multidisciplinary meeting (MDM) to perform surgery following identification of residual disease on PET–CT. As for planned NDs, both modified radical ND and selective NDs were acceptable, provided that the conditions were met. The type of access incision was left to the surgeon to decide.

Recommended clinical guidelines for PET–CT scanning

1. Patient scheduling:

    - PET–CT should have ideally been performed between 10 and 12 weeks after completion of the last dose of CRT. However, it could be performed between 9 and 13 weeks after completion of CRT and before any biopsies.

2. Preparation:

    - Patients had to fast for 6 hours prior to the scan.
    - Patients had to be weighed without shoes and coats (ensuring that the weighing device was calibrated).
    - Blood glucose had to be recorded using Boehringer Mannheim's Glucometer (Boehringer Mannheim, Mannheim, Germany) (ensuring that the device was calibrated).
    - Patients had to drink 2–3 glasses of water prior to the test to ensure hydration.
    - Metal denture fixtures had to be removed whenever possible (to reduce CT artefacts and improve semiquantitative accuracy).

3. Injection:
    - Fludeoxyglucose had to be injected via a butterfly cannula under quiet conditions.
    - The patient had to be asked to remain silent.
    - The injected dose had to be 4.5 MBq of fludeoxyglucose per kilogram of body weight up to a maximum of 400 MBq.
4. Fludeoxyglucose uptake:
    - The patient had to remain inactive in a comfortable and quiet environment during the uptake.
    - The patient had to empty their bladder just prior to positioning on scanner bed.
    - The emission scan had to start at 90 minutes post injection.
5. Positioning:
    - The patient had to be scanned on a regular couch top.
    - Scanning had to begin at the skull vertex and end at the groin.
    - The scan had to be done with arms down if a single whole-body scan was performed. If the body was scanned separately from the H&N, then the body had to be scanned with the arms up.
6. Acquisition parameters:
    - Each PET–CT centre was advised to use routine local protocols.
7. Reconstruction parameters:
    - The PET–CT centre was advised to use ordered subset expectation maximisation, with CT for attenuation correction using local routine parameters.
8. Archive:
    - The reconstructed CT, PET attenuation-corrected and non-attenuation-corrected data had to be archived locally.
9. QA and quality control:
    - This was done under an agreed QA and quality control protocol.
10. Reporting PET–CT findings:
    - A standardised criterion was used for reporting PET–CT findings.
    - The PET–CT findings had to be reported by the local hospital imaging team.
    - The PET–CT findings had to be reviewed at the central laboratory by an expert in the PET core laboratory, who was independent of the local report.
    - Differences in reporting between the local hospital and the central laboratory were resolved by consensus, otherwise a third expert at the core laboratory reported on the scan and the majority view was taken.
11. PET–CT data transfer to the central laboratory:
    This was under an agreed protocol transfer following reconstructed and anonymised files:
    - CT
    - PET attenuation corrected
    - PET non-attenuation corrected
    - PET–CT report from local imaging team.

12. Adverse events and serious adverse events (SAEs) during PET CT:

    - Adverse events and SAEs were required to be recorded and documented by the local investigational team in accordance with agreed protocols. However, none of these was reported.

Quality assurance review of PET–CT scans

For patients in the PET–CT arm of the trial, the PET–CT scan was reviewed by both the local reporter and the trial core laboratory radiologist. When there were significant differences between the reports of the local PET–CT clinician and the central PET–CT reviewer that might have affected patient management, the central PET–CT reviewer (from the core laboratory) and the participating centre radiologist conferred and a final report was agreed. If agreement could not be reached, a second independent reviewer from the core laboratory was asked to review the scans and issue a second opinion report. If there was still a significant difference between the local PET–CT report and the core laboratory reports, the final decision on which report to use rested with the local PET–CT clinician, local treating clinicians and MDT.

Any discordance between the reports of the local PET–CT clinician and central PET–CT reviewer was documented in the patient's assessment form including which report was used to make the final clinical decision.

Quality assurance review of staging and response scans

Central QA review was performed on 10% of staging scans (and response scans for patients in the ND arm) to check for concordance. The results of the review are not yet finalised and this will be reported in a later paper.

Patient assessments

The schedule of assessments is shown in *Table 2*. All biobank samples collected at baseline and follow-up will be retained for future work.

Patient follow-up

Patient follow-up was monthly for the first year up to the 12th month after randomisation, and 2-monthly for the second year up to the 24th month after randomisation. It was recommended, but not mandatory, that patients should have chest radiography annually.

At each follow-up visit, the patient received a full clinical examination with palpation of the neck and visualisation of primary site when possible, with appropriate trial forms and questionnaires completed at the completion of CRT, and at 6, 12 and 24 months after randomisation.

Patients with confirmed recurrence

Recurrent disease was defined as a confirmed tumour in the primary site after a clear 3-month post-CRT assessment or metastatic nodal disease identified in the neck after negative PET–CT post CRT or after a ND, during the follow-up period. A tumour was confirmed as recurrent if disease was confirmed by biopsy or radiological cross-sectional imaging.

Patients with recurrence, confirmed by biopsy or needle aspiration and CT/MRI or by their PET–CT scan were referred back to the treating clinician and MDM urgently for consideration for salvage treatment, including ND for patients with isolated nodal recurrence.

Serious adverse events

Investigators were required to inform the trials unit immediately of any SAEs following CRT, PET–CT or ND.

TABLE 2 Schedule of investigation

Assessments	Time point								
	Pre-study entry	2 weeks post ND[a]	Last CRT dose	Monthly follow-up visits during year 1	Post-CRT assessment[b]	6-month follow-up[b]	12-month follow-up[b]	2-monthly follow-up visits during year 2	24-month follow-up[b]
MDM discussion	✗								
Informed consent for trial	✗								
Evaluations									
Complete history/physical	✗			✗[c]	✗	✗[c]	✗[c]	✗[c]	✗[c]
Complications reporting		✗[d]			✗[e]				
Examination with or without biopsy[f]	✗[g]		✗[h]		✗				
Imaging									
H&N CT/MRI	✗[i]				✗[i]				
Thoracic CT	✗[g]								
PET–CT	(✗[i])				✗[k,l]				
CXR							✗		✗

continued

TABLE 2 Schedule of investigation (*continued*)

Assessments	Time point								
	Pre-study entry	2 weeks post ND[a]	Last CRT dose	Monthly follow-up visits during year 1	Post-CRT assessment[b]	6-month follow-up[b]	12-month follow-up[b]	2-monthly follow-up visits during year 2	24-month follow-up[b]
Quality-of-life assessments									
EORTC QLQ-C30 and H&N35	✗		✗[m]			✗	✗		✗
MDADI	✗		✗[m]			✗	✗		✗
EQ-5D	✗		✗[m]			✗	✗		✗
Health economic questionnaires[n]	✗		✗[m]		✗	✗	✗		✗
Blood sample collection[o]	✗				✗[o]		✗[o]		

CXR, chest radiography; EORTC, European Organisation for Research and Treatment of Cancer; EQ-5D, EuroQoL-5 Dimensions; H&N35, head and neck module with 35 questions; MDADI, MD Anderson Dysphagia Inventory; QLQ-C30, Quality of Life Questionnaire for Cancer (with 30 questions).

a For patients having a planned ND after CRT completion, if this fell within 2 weeks of, or coincided with, the 3-month post-CRT assessment, only assessments stated under the 3-month post-CRT assessment had to be performed.
b Required to be performed within 2 weeks of the assessment time stated.
c If recurrence was confirmed on investigation, the lead investigator was required to complete the recurrence form.
d Complications following ND were assessed at the follow-up visit at 2 weeks post ND, and reported on the surgery form.
e Complications following CRT were assessed and documented at the post-CRT assessment, and reported using the CRT forms.
f Examination was ideally under anaesthetic; however, a clinical examination was also acceptable.
g All pre-randomisation staging assessments were ideally done within 4 weeks of randomisation or at least the last staging test was required to be within 4 weeks of randomisation.
h Patients having a ND after CRT were required to have a clinical examination, CT with or without EUA before ND and after CRT.
i Required to be done before EUA or at least within 2 weeks.
j Patients with occult primary disease only were required to have a diagnostic PET–CT pre randomisation.
k Ideally done at 10–12 weeks after completion of CRT; however, the allowed range was between 9 and 13 weeks.
l PET–CT done for experimental group patients was required to be done before EUA.
m Required to be completed within 2 weeks of completing CRT.
n Only in sites that were taking part in the health economics evaluation.
o 20-ml blood samples were collected at baseline using containers provided by the PET-NECK COLLECT trial office and processed in accordance with to sample collection protocol. For the post-CRT assessment bloods, these were required to be taken at 3 months post CRT completion (or within 2 weeks of these times). For the 12-month follow-up bloods, these need to be taken at 12 months post CRT completion (or within 2 weeks of these times).

Each time the patient was seen in clinic he or she was asked if any SAEs had occurred. The occurrence of SAEs was based on information provided by either patients or their carers. The following adverse events were considered serious:

- death
- life-threatening disease
- hospitalisation or prolongation of hospitalisation
- congenital abnormality
- persistent disability
- other medically significant event.

Trial-specific exclusions from reporting SAEs included any of the following:

- elective hospitalisations for the treatment of the primary cancer and its effects
- elective hospitalisation for social reasons
- elective hospitalisation for pre-existing conditions that were not exacerbated by the treatment.

Adverse effects of PET–CT scanning

If, after administration of the radiopharmaceutical required for PET–CT, the biodistribution within the patient was not as would be expected, the ARSAC certificate holder had to be informed and a check made to determine that the correct radiopharmaceutical was administered. If the correct radiopharmaceutical was administered, the supplier of the radiopharmaceutical had to be informed.

In the case of an incorrect radiopharmaceutical being administered to a patient, the ARSAC licence certificate holder and a senior member of staff had to be informed immediately who was then required to follow the guidance in Health and Safety Guidance Number 95.[55]

(a) The ARSAC certificate holder was required to inform the patient of the error.
(b) The referring clinician had to be informed.
(c) The GP of the patient had to be informed.
(d) A risk assessment was required to be requested from the radiation protection advisor.
(e) The staff member was required to inform the chief investigator and the sponsor.

A report of any incident, including action taken in order to prevent a recurrence, was required to be compiled by the radiation protection supervisor and forwarded to the radiation protection advisor. Any such incident was required to be recorded as an unexpected SAE and also reported through the individual trust critical incident policies.

Patient withdrawal

Patients were informed about the right to withdraw from the study at any time without giving a reason. However, site staff were asked to make every effort to identify the reason for withdrawal. Withdrawn patients were asked if any data collected prior to their withdrawal could be used in the analyses.

The reason for withdrawal (when known) was recorded on the patient notes and the trial office was informed of the withdrawal.

Patients who elected to withdraw from the trial interventions remained in the trial for follow-up per protocol unless they withdrew their consent.

Patients who changed their mind about withdrawal and wished to rejoin the study could chose to do so at any time, but they needed to be reconsented and follow-up data were to be collected only from that point onwards.

Patients moving out of the area

When patients moved away from the area to another participating centre, every effort had to be made, with the patients' consent, to transfer the follow-up of patients so that the new centre could take over the responsibility for follow-up. Close co-ordination with the PET-NECK trial office was essential for this.

Outcomes

Overall survival (primary outcome)

Information on death and survival was obtained from centres via death and follow-up forms.

Cost-effectiveness (primary outcome)

Data for the health economic analysis were collected from the trial EuroQoL-5 Dimensions (EQ-5D) and resource use questionnaires, as well as the case report forms. Full details of the health economic analysis are given in *Chapter 4*.

Complication rates

Complications following ND surgery were reported.

Disease-specific survival

Causes of death were reported on the death form. Deaths were regarded as caused by H&N cancer if they were reported as such, if there was persistent, recurrent or metastatic disease present at death or if the patient died as a result of complications of treatment for H&N cancer.

Recurrence in neck nodes

Details of recurrences were reported at follow-up visits. Any notification of recurrence within 3 months of radiotherapy was regarded as persistent disease and notifications after that date were regarded as recurrences. Recurrences in the neck nodes were reported as ipsilateral or contralateral in the notification of recurrence.

Quality of life

The questionnaires included the following instruments:

- EQ-5D: five 3-point scales and one summary 100-point scale. This was used for the health economics evaluation
- European Organisation for Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire for Cancer (with 30 questions) (QLQ-C30): five functional, three symptom and a global scale and six single items for assessment of general quality of life
- EORTC Quality of Life Questionnaire for Cancer head and neck module with 35 questions (QLQ-H&N35): seven scales and 11 single items for H&N cancer-related quality of life
- MD Anderson Dysphagia Inventory (MDADI): an overall function scale and four subscales.

Patients were requested to complete these questionnaires at five time points. The first was completed in clinic prior to randomisation. Later questionnaires were received either in clinic or through the post before the assessment dates were due at these time points: within 2 weeks of CRT completion, 6 months after randomisation, and 12 and 24 months after randomisation.

Accuracy of PET–CT scanning

Positron emission tomography–computerised tomography scans were assessed both by the local radiologist and by the trial core radiologist.

Sample size

The trial was designed to allow demonstration of non-inferiority of the PET–CT surveillance arm compared with the control arm (planned ND), which was assumed to have a 2-year OS of 75%. The margin for non-inferiority was set at 10%, implying that the 2-year OS of the experimental arm should not be below 65%, a difference equivalent to a HR of 1.50.

The trial was expected to recruit for 3 years, with an additional follow-up period of 2 years. The sample size was set at 560 patients in total (280 in each treatment arm), giving a 90% power with a type I error of 5%, allowing for a 3% loss to follow-up. The sample size was calculated by simulation assuming unadjusted analysis by a Cox's proportional hazards model and assuming follow-up to the end of the trial 5-year period (using data from the Office for National Statistics). If no follow-up existed beyond the 2-year requested data, then the power would be 76%. It was expected that the true power would lie between these figures (i.e. 76% and 90%). The assumptions underlying the sample size calculation were monitored by the Independent Data and Safety Monitoring Committee (IDSMC) throughout the study.

Randomisation

Treatments were centrally allocated through the PET-NECK trial office at the Warwick Clinical Trials Unit. Allocation was performed using a minimisation algorithm balancing by the following: centre, timing of planned ND (before or after CRT), chemotherapy schedule (concomitant platinum, concomitant cetuximab, neoadjuvant and concomitant platinum, neoadjuvant TPF with concomitant platinum, other approved), disease site (oropharyngeal, laryngeal, oral, hypopharyngeal or occult), T stage (T1–T2 vs. T3–T4) and N stage (N2a–N2b vs. N2c–N3). At randomisation, the eligibility of the patient was checked and the trial number and treatment arm were allocated. Each patient was asked to complete a quality-of-life questionnaire (before being informed of the randomisation allocation).

Allocation concealment

When the randomisation service was telephoned, participant details were taken and registration into the trial followed. The allocation was then generated, ensuring concealment.

Blinding

It was not possible to blind either the doctor or the patient to the allocated treatment.

Statistical methods

Participants were analysed according to the treatment group to which they were randomised. Analyses were guided by an analysis plan prepared before data were available.

A preplanned early stopping guideline was applied, assuming that at least 2 years' follow-up would be available for each patient, and that 140 deaths (25%) would be expected. Three analyses of the primary outcome were prespecified, equally spaced at 47, 94 and 140 deaths, to be viewed only by the IDSMC and trial statistician. At each of these points, it could be concluded that the experimental treatment is equivalent (non-inferior). The 5% one-sided type I error rate for testing non-inferiority was controlled by an O'Brien–Fleming-like alpha spending rule set at *p*-values of 0.004, 0.007 and 0.047. It could alternatively be concluded that the experimental arm is inferior. A similar 10% type I rate for testing non-equivalence was used, with *p*-values set at 0.02, 0.032 and 0.084. At the first interim analysis (48 deaths), when

84% of patients had been randomised, the *p*-value for non-inferiority was 0.0025, which was below the (0.004) threshold for consideration of stopping. The IDSMC considered that data were immature and that the trial should continue. At the second interim analysis (102 deaths) the *p*-value for non-inferiority was 0.008, that is, above the threshold for the second analysis, and the trial continued to the final analysis.

Demographic and clinical characteristics and baseline measurements are presented to evaluate the comparability of intervention groups and generalisability to clinical settings. A Consolidated Standards of Reporting Trials (CONSORT) flow diagram was produced.[56]

The primary end point, OS, was measured for every patient from the date of randomisation to the date of death. Survival times for patients on follow-up and patients lost to follow-up were censored at the latest date at which they were known to be alive. Kaplan–Meier survival curves were plotted. To test non-inferiority of the surveillance arm the HR was estimated using a Cox's proportional hazards model stratified by intended timing of planned ND (before or after CRT), with the trial treatment arm as the only covariate. Non-inferiority is demonstrated by the 95% quantile of the estimated HR of the surveillance arm being less than the inferiority limit (HR 1.50). Proportionality of hazards was checked using a time-dependent treatment effect. A secondary analysis adjusting for N stage, T stage, tumour site, chemotherapy schedule and timing, sex and centre was performed. Kaplan–Meier OS was also calculated for the planned timing of ND (before or after CRT) subgroups and also for p16-positive and -negative tumours. The treatment effect on OS was also presented for subgroups of site of tumour, T stage, planned chemotherapy schedule and p16 status, using HR plots with interaction statistics using methods described by the Early Breast Cancer Trialists' Collaborative Group.[57]

Causes of death were classified as H&N cancer or other causes. Cumulative incidence rates by treatment arm were plotted for these competing risk groups and the difference between treatments was tested using Gray's test.[58]

Patients were deemed at risk of recurrence when they completed their CRT and time to recurrence was measured from that date. This meant that patients in the planned ND group who underwent surgery before CRT were not at risk until they had finished all their treatment. Counts of neck node recurrences by treatment were reported.

Quality-of-life scales at five time points were available for the EORTC QLQ-C30 version 3, EORTC QLQ-H&N35 H&N-specific quality-of-life questionnaire and the MDADI. The recommended scoring methods were applied[59] (see also Fayers[60] and Chen *et al.*[61]). The differences between treatments in mean changes from baseline scores were calculated and compared for each of the assessment points separately using the Wilcoxon signed-rank test. The proportion of patients whose scores changed by at least 10% was also reported.

Complications of ND surgery were reported as counts and as proportion of NDs with Agresti–Coull binomial CIs.

Positron emission tomography–computerised tomography accuracy was difficult to assess, as it is difficult to define a false positive. Therefore, it was not possible to measure PET–CT accuracy in a meaningful way.

Concordance of local PET–CT scan assessments with that of the trial review radiologist were reported.

Statistical analyses were performed using SAS version 9.3 software (SAS Institute Inc., Cary, NC, USA).

Database and data processing

The database was held on the Warwick Clinical Trials Unit's Microsoft Structured Query Language server system (Microsoft Corporation, Redmond, WA, USA) and imposed rules for data entry that included valid range for responses, linked dates and patient identification numbers.

Data were single entered into the database by study personnel. The trial statistician carried out checks of plausibility of values, missing data and form return rates to enable further queries to be resolved prior to freezing data for scheduled IDSMC reports and analysis. A 100% check on the details in death reports was applied.

Chapter 3 Results

Screening and recruitment

Recruitment

A total of 564 patients were recruited between 2 October 2007 and 23 August 2012: 282 to each trial arm (*Figure 2*).

Two patients in the surveillance arm were subsequently found to be ineligible: one patient's disease was found to be T1N1 and another patient had metastatic disease at the time of randomisation. Follow-up data were collected and the patients are included in analyses.

The participant flow diagram outlines the passage of patients through the study; more detail is given in later sections (*Figure 3*).

Screening

All patients newly diagnosed with HNSCC were considered potentially eligible and had to be screened by the MDT prior to their clinic appointment.

Both screened patients and those approached for study participation had to be recorded anonymously on the screening log. The log was updated monthly and passed to the co-ordinating centre.

If a screened patient was not eligible for the trial, an anonymous record of the case was recorded on the screening log. Recorded details included the name of the centre, the patient's initials and the reason for non-randomisation, if not randomised. Patients randomised to the trial were also recorded.

In total, 1792 patients were screened. The number of patients screened by centre is given in *Table 3*.

Of the 1792 patients screened, 564 were randomised, 1032 did not fulfil the eligibility criteria, and 196 patients refused consent. *Table 4* shows reasons for patients not meeting the standard eligibility criteria and *Table 5* shows reasons for unwillingness of eligible patients to enter the study.

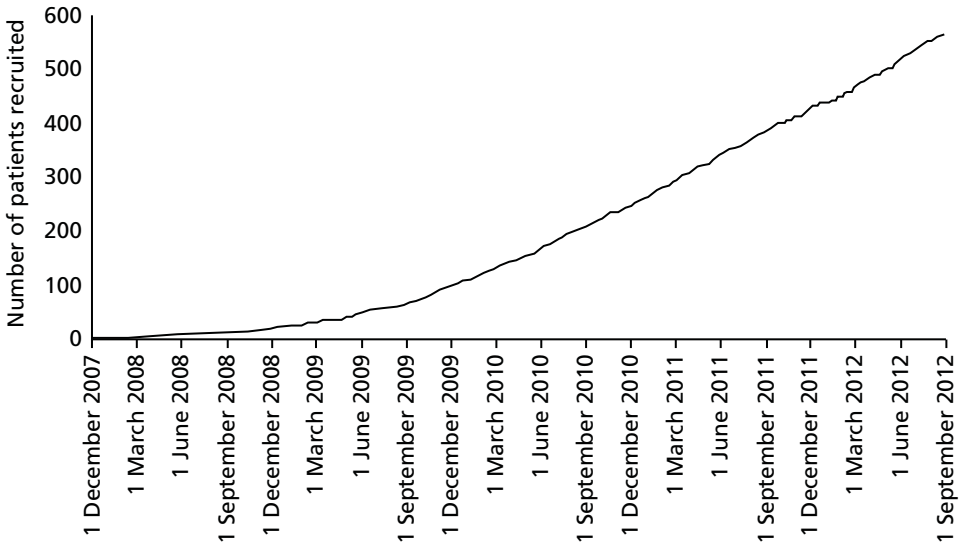


FIGURE 2 Recruitment.

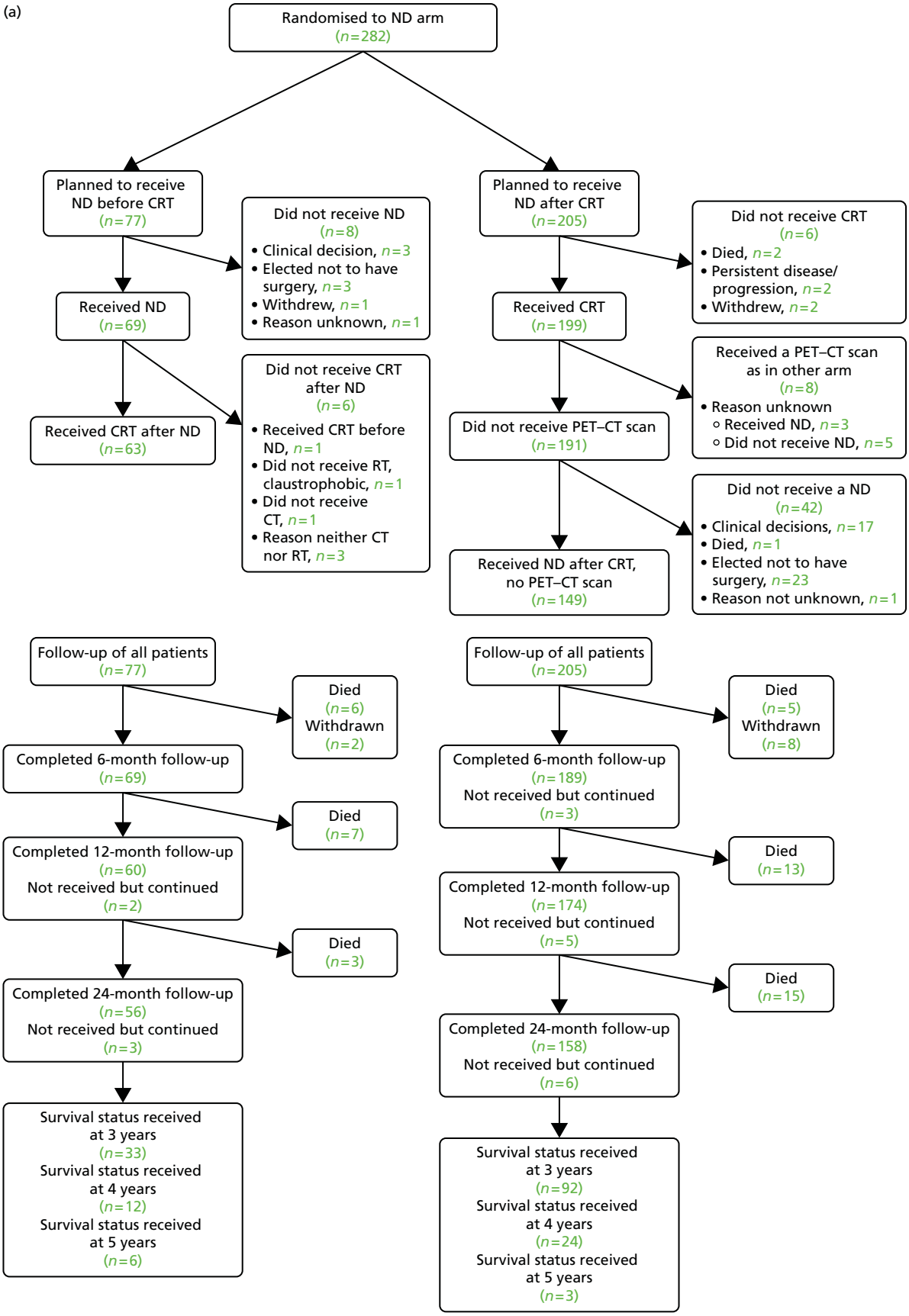


FIGURE 3 Participant flow diagram. (a) ND arm; and (b) surveillance arm. Adapted from the New England Journal of Medicine, Mehanna H, Wong W-L, McConkey CC, Rahman JK, Robinson M, Hartley AGJ, PET-CT Surveillance versus Neck Dissection in Advanced Head and Neck Cancer, 374, 1444–54. Copyright © (2016) Massachusetts Medical Society. Reprinted with permission.[62] (*continued*)

(b)

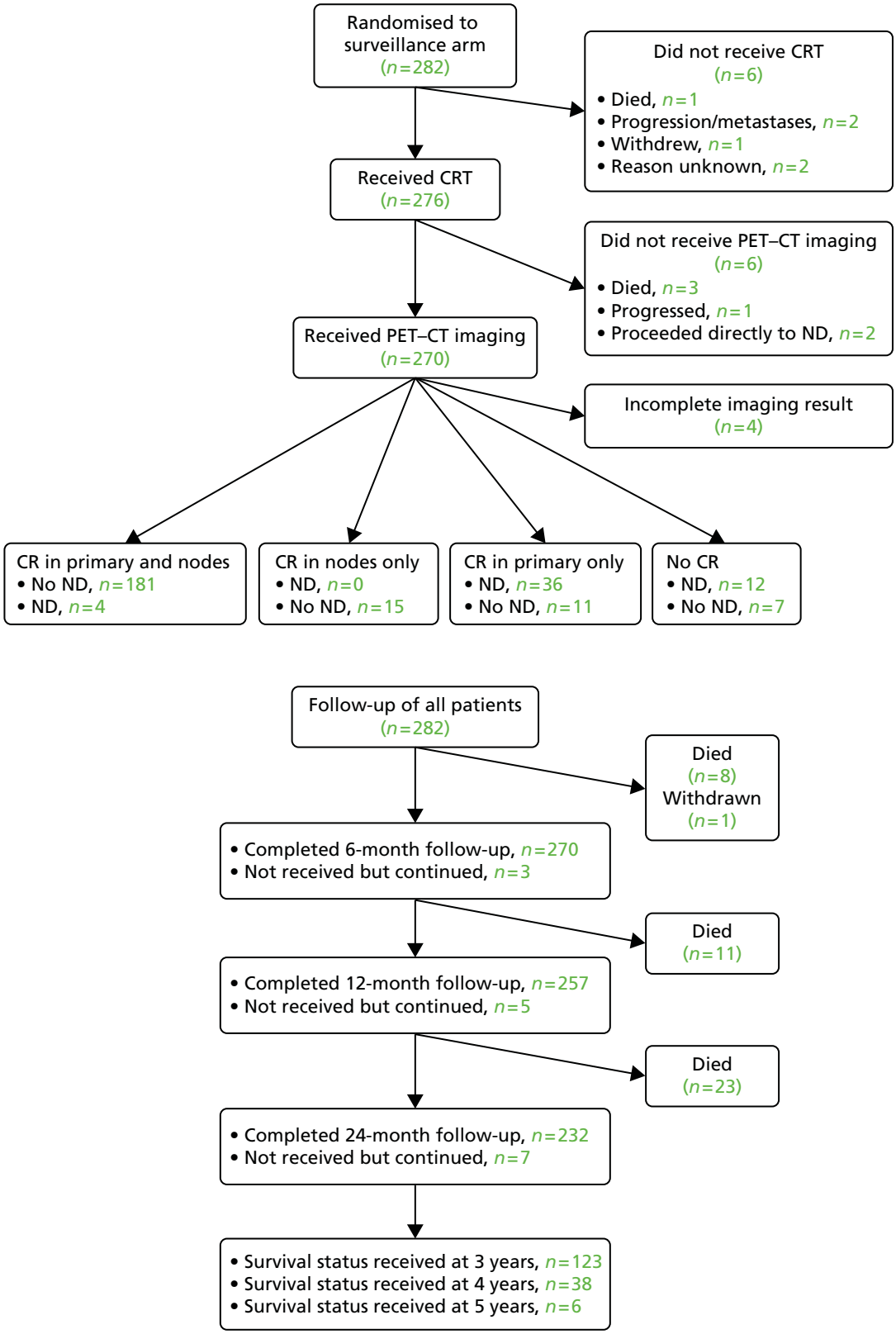


FIGURE 3 Participant flow diagram. (a) ND arm; and (b) surveillance arm. Adapted from the New England Journal of Medicine, Mehanna H, Wong W-L, McConkey CC, Rahman JK, Robinson M, Hartley AGJ, PET-CT Surveillance versus Neck Dissection in Advanced Head and Neck Cancer, 374, 1444–54. Copyright © (2016) Massachusetts Medical Society. Reprinted with permission.[62]

TABLE 3 Number of patients screened by centre

Site	Number of patients screened
Aberdeen Royal Infirmary	7
Barnet and Chase Farm Hospitals	2
Beatson West of Scotland Cancer Centre	31
Belfast City Hospital	33
Blackpool Victoria Hospital	29
Birmingham Heartlands Hospital	70
Bradford Royal Infirmary	23
Bristol Haematology and Oncology Centre	40
Castle Hill Hospital	22
Cheltenham General Hospital	5
Christie Hospital	25
Derriford Hospital	12
Guy's and St Thomas' Hospitals	41
James Cook University Hospital	10
Kent and Canterbury Hospital	77
Mount Vernon Hospital	17
New Cross Hospital	11
North Manchester General Hospital	12
Nottingham University Hospital	17
Poole Hospital	10
Queen Alexandra Hospital	5
Queen Elizabeth Hospital Birmingham	77
Royal Blackburn Hospital	126
Royal Derby Hospital	59
Royal Devon and Exeter Hospital	35
Royal Marsden Hospital	5
Royal Preston Hospital	21
Royal Surrey County Hospital	42
Royal United Hospital	44
Russells Hall Hospital	11
Singleton Hospital	101
Sir Bobby Robson Cancer Trials Research Centre (Freeman Hospital)	32
Southend University Hospital	4
Sunderland Royal Hospital	153
Torbay Hospital	11
UHCW	143
University College London Hospital	32
University Hospital Aintree	131

TABLE 3 Number of patients screened by centre (*continued*)

Site	Number of patients screened
Velindre Cancer Centre	26
Walsall Manor Hospital	12
Western General Hospital	39
Weston Park Hospital	188
William Harvey Hospital	1

TABLE 4 Known reasons for ineligibility of screened patients

Ineligibility criteria	Number of patients in category
Patient did not have the required histological diagnosis	187
Patient did not have N2 or N3 disease	223
Patient had distant metastases	57
Patient had previous treatment for HNSCC	59
Previous cancer in past 5 years	4
Patient would not receive protocol-driven CRT regimen	26
Patient was unfit for surgery and/or CRT	115
Patient required resection of tumour	151
Inoperable nodes	5
Patient required palliative treatment	69
For radiotherapy only	3
Clinical decision not to enter patient in the trial	15
Other treatment planned	41
Other reasons not specified above	47
Reason for ineligibility not specified	30
Total	1032

TABLE 5 Reasons for unwillingness of eligible patients to participate

Reason	Number of patients in category
Patient wanted ND or standard treatment	47
Patient did not want surgery	36
Patient wanted more control over the type or place of treatment	16
Patient had anxieties about the large amount of information and making decisions on randomisation and treatment	12
Patient did not want to take part in clinical study	11
Total	122

The majority of patients who were ineligible either had the wrong type or severity of disease or a different type of treatment was indicated. There were 47 patients who were not approached about the study for reasons such as cognitive impairment, psychiatric issues or short-term memory problems that were likely to affect compliance. Other examples include language barriers, too much information to give to a patient as a result of having to tell them about their disease at the same time as giving study information, lack of clinicians available to discuss trial/consent patients, patient travelling elsewhere for treatment, a patient refusing to be treated or a patient being entered into another incompatible study. For 30 patients, the reason for ineligibility was not known.

Of the 196 patients who refused to participate, 74 did not give a reason for non-participation. The reasons for non-participation of the other 122 patients are given below.

One patient was anxious about being randomised in the trial, 11 patients were not interested in participating in a clinical trial, two patients did not want to travel for PET–CT if randomised to the surveillance arm of the trial, one patient felt that more imaging would be too much, four patients were overwhelmed with the amount of information given (trial information as well as diagnosis and treatment information), seven patients were too distressed after receiving diagnosis information to make a decision about the trial, one patient requested treatment at another clinic, one patient wanted to remain private, one patient wanted to retain control over his/her treatment and 93 patients had specific treatment preferences. Of these, 42 patients preferred ND and 36 patients did not want ND. There were two patients whose treatment preference was not specified. One patient wanted radiotherapy only, one wanted photodynamic therapy, four wanted CRT only, one wanted CRT followed by PET and one patient wanted to proceed immediately to adjuvant chemotherapy. Finally, five patients wanted standard CRT with ND, before or after CRT.

Baseline characteristics

Baseline characteristics of participants by trial arm

Treatment allocation by minimisation was balanced by the six characteristics in *Table 6*. A total of 38 centres took part, randomising mainly oropharyngeal patients (84.4%). Neoadjuvant chemotherapy was intended in 35.8% of patients.

TABLE 6 Baseline characteristics balanced at treatment allocation

Baseline characteristic	Trial arm, *n* (%)	
	Surveillance	ND
Centre		
Arden	30 (10.6)	30 (10.6)
Sheffield and Chesterfield	25 (8.9)	25 (8.9)
Birmingham	23 (8.2)	24 (8.5)
Bristol	19 (6.7)	19 (6.7)
Edinburgh	17 (6.0)	19 (6.7)
Lancashire	14 (5.0)	14 (5.0)
Glasgow	13 (4.6)	12 (4.3)
Royal Surrey	12 (4.3)	13 (4.6)
East Peninsula	11 (3.9)	11 (3.9)
Guy's and St Thomas'	11 (3.9)	11 (3.9)
Newcastle	11 (3.9)	11 (3.9)
Black Country	10 (3.5)	9 (3.2)
Manchester	11 (3.9)	8 (2.8)

TABLE 6 Baseline characteristics balanced at treatment allocation (*continued*)

Baseline characteristic	Trial arm, *n* (%)	
	Surveillance	ND
Cardiff	9 (3.2)	9 (3.2)
Swansea	8 (2.8)	9 (3.2)
East and North Herts	6 (2.1)	7 (2.5)
Bath	5 (1.8)	7 (2.5)
West Peninsula	6 (2.1)	5 (1.8)
Nottingham	5 (1.8)	5 (1.8)
Sunderland	4 (1.4)	6 (2.1)
Derby	5 (1.8)	3 (1.1)
East Kent	5 (1.8)	3 (1.1)
Liverpool	3 (1.1)	4 (1.4)
East Lancashire	3 (1.1)	2 (0.7)
Gloucestershire	3 (1.1)	2 (0.7)
South Tees	3 (1.1)	2 (0.7)
Hull & East Yorkshire	1 (0.4)	3 (1.1)
Portsmouth	2 (0.7)	2 (0.7)
Dorset	1 (0.4)	2 (0.7)
Essex	3 (1.1)	–
Barnet	1 (0.4)	1 (0.4)
Aberdeen	2 (0.7)	–
Bradford	–	1 (0.4)
Belfast	–	1 (0.4)
Royal Marsden	–	1 (0.4)
University College London Hospital	–	1 (0.4)
ND policy before or after CRT		
ND before CRT	76 (27.0)	77 (27.3)
ND after CRT	206 (73.1)	205 (72.7)
Approved chemotherapy schedules		
Concomitant platinum	163 (57.8)	169 (59.9)
Concomitant cetuximab	14 (5.0)	14 (5.0)
Neoadjuvant and concomitant platinum	10 (3.6)	9 (3.2)
Neoadjuvant TPF with concomitant platinum	89 (31.6)	85 (30.1)
Other agreed schedules	6 (2.1)	5 (1.8)
Tumour site		
Oral	4 (1.4)	7 (2.5)
Oropharyngeal	240 (85.1)	236 (83.7)
Laryngeal	18 (6.4)	19 (6.7)
Hypopharyngeal	15 (5.3)	14 (5.0)
Occult H&N	5 (1.8)	6 (2.1)

continued

TABLE 6 Baseline characteristics balanced at treatment allocation (*continued*)

Baseline characteristic	Trial arm, *n* (%)	
	Surveillance	ND
T stage		
T1/T2	162 (57.4)	160 (56.7)
T3/T4	116 (41.1)	116 (41.1)
Occult H&N	4 (1.4)	6 (2.1)
N stage		
N2a/N2b	221 (78.4)	222 (78.7)
N2c/N3	61 (21.6)	60 (21.3)

Adapted from the New England Journal of Medicine, Mehanna H, Wong W-L, McConkey CC, Rahman JK, Robinson M, Hartley AGJ, PET-CT Surveillance versus Neck Dissection in Advanced Head and Neck Cancer, 374, 1444–54. Copyright © (2016) Massachusetts Medical Society. Reprinted with permission.[62]

Other baseline characteristics of participants by trial arm

The patient characteristics in *Table 7* were collected either at randomisation or via baseline information forms. The mean age was 57.9 years and 83.3% were male. Three-quarters of participants were either current or past smokers and, of those tested for p16 status, 335 out of 446 (75%) were positive.

Further details of the tumour and nodal stages of randomised patients are given in *Tables 8* and *9*.

TABLE 7 Further baseline characteristics by trial arm

Baseline characteristic	Trial arm	
	Surveillance	ND
Age (years)		
n	282	282
Mean (SD)	57.6 (7.5)	58.2 (8.1)
Median (IQR)	57 (53–63)	58 (53–63)
Minimum, maximum	37, 79	38, 83
Sex,* n *(%)		
n	282	282
Male	223 (79.1)	237 (84.0)
Female	59 (20.9)	45 (16.0)
Primary site,* n *(%)		
n	277	277
Tonsil	138 (49.1)	134 (48.4)
Base of tongue	82 (29.6)	83 (30.0)
Floor of mouth	1 (0.4)	–
Palate	4 (1.4)	3 (1.1)
Tongue	1 (1.1)	2 (0.4)
Supraglottis	15 (5.4)	17 (6.1)
Glottis/subglottis/transglottis	2 (0.7)	1 (0.7)

TABLE 7 Further baseline characteristics by trial arm (*continued*)

Baseline characteristic	Trial arm	
	Surveillance	ND
Pyriform fossa	14 (5.1)	12 (4.3)
Posterior pharyngeal wall	3 (1.1)	6 (2.2)
More than one site	17 (6.1)	19 (6.9)
Tonsil, base of tongue	9	9
Tonsil, base of tongue, pyriform fossa	1	–
Tonsil, base of tongue, palate	1	–
Tonsil, base of tongue, post-pharyngeal wall	–	1
Tonsil, palate, tongue	–	1
Tonsil, palate	1	2
Tonsil, tongue	–	1
Tonsil, pyriform fossa	–	1
Tonsil, supraglottis, pyriform fossa	–	1
Tonsil, posterior pharyngeal wall	1	–
Tonsil, supraglottis	1	–
Base of tongue, supraglottis	–	2
Base of tongue, floor of mouth	1	1
Base of tongue, posterior pharyngeal wall	1	–
Supraglottis, posterior pharyngeal wall	1	–
T stage,* n *(%)		
n	282	282
T1	48 (17.0)	52 (18.4)
T2	114 (40.4)	108 (38.3)
T3	61 (21.6)	52 (18.4)
T4	55 (19.5)	64 (22.7)
Occult	4 (1.4)	6 (2.1)
N stage,* n *(%)		
n	282	282
N2a	54 (19.1)	44 (15.6)
N2b	167 (59.2)	178 (63.1)
N2c	52 (18.4)	52 (18.4)
N3	9 (3.2)	8 (2.8)
Side of primary,* n *(%)		
n	280	278
Left	120 (42.9)	102 (36.7)
Right	129 (46.1)	142 (51.1)
Midline and/or left and right	31 (11.1)	34 (12.2)

continued

TABLE 7 Further baseline characteristics by trial arm (*continued*)

Baseline characteristic	Trial arm	
	Surveillance	ND
Differentiation,* n *(%)		
n	237	235
Well differentiated	15 (6.3)	10 (4.3)
Moderately differentiated	101 (42.6)	90 (38.3)
Poorly differentiated	119 (50.2)	134 (57.0)
Undifferentiated	2 (0.8)	1 (0.4)
Type of staging scan,* n *(%)		
n	282	281
PET–CT	27 (9.6)	27 (9.6)
CT	169 (59.9)	173 (61.6)
MRI	86 (30.5)	81 (28.8)
ECOG performance status,* n *(%)		
n	281	281
0	220 (78.3)	218 (77.6)
1	60 (21.4)	60 (21.4)
2	1 (0.4)	3 (1.1)
Smoking,* n *(%)		
n	281	281
Current	88 (31.3)	76 (27.0)
Past	119 (42.3)	134 (47.7)
Never	74 (26.3)	71 (25.3)
Alcohol,* n *(%)		
n	280	279
Current	222 (79.3)	234 (83.9)
Past	34 (12.1)	18 (6.5)
Never	24 (8.6)	27 (9.7)
Ethnic group,* n *(%)		
n	281	280
White	280 (99.6)	278 (99.3)
Black or black British	1 (0.4)	2 (0.7)
p16 status,* n *(%)		
n	226	220
p16 positive	164 (72.6)	171 (77.7)
p16 negative	62 (27.4)	49 (22.3)
p16 test not done or result not available	56	62

ECOG, Eastern Cooperative Oncology Group; IQR, interquartile range; SD, standard deviation.
Adapted from the New England Journal of Medicine, Mehanna H, Wong W-L, McConkey CC, Rahman JK, Robinson M, Hartley AGJ, PET-CT Surveillance versus Neck Dissection in Advanced Head and Neck Cancer, 374, 1444–54. Copyright © (2016) Massachusetts Medical Society. Reprinted with permission.[62]

TABLE 8 The TNM stages by trial arm (1)

	ND intended (*n*)			
	Before CRT		After CRT	
Stage	Surveillance	ND	Surveillance	ND
T1N2a	3	4	12	11
T1N2b	5	12	24	21
T1N2c	1	2	3	–
T1N3	–	1	–	1
T2N2a	14	4	12	11
T2N2b	21	24	48	50
T2N2c	2	4	12	12
T2N3	1	1	4	2
T3N2a	2	1	7	5
T3N2b	11	5	24	25
T3N2c	3	5	11	9
T3N3	1	1	2	1
T4N2a	–	4	2	4
T4N2b	8	4	25	32
T4N2c	3	2	17	17
T4N3	–	–	–	1
Occult N2a	–	–	2	–
Occult N2b	1	3	–	2
Occult N2c	–	–	–	1
Occult N3	–	–	1	–

TABLE 9 The TNM stages by trial arm (2)

	ND intended, *n* (%)			
	Before CRT		After CRT	
Stage	Surveillance	ND	Surveillance	ND
T stage				
T1	9 (11.8)	19 (24.7)	39 (18.9)	33 (16.1)
T2	38 (50.0)	33 (42.9)	76 (36.9)	75 (36.6)
T3	17 (22.4)	12 (15.6)	44 (21.4)	40 (19.5)
T4	11 (14.5)	10 (13.0)	44 (21.4)	54 (26.3)
Occult	1 (1.3)	3 (3.9)	3 (1.5)	3 (1.5)
N stage				
N2a	19 (25.0)	13 (16.9)	35 (17.0)	31 (15.1)
N2b	46 (60.5)	48 (62.3)	121 (58.7)	130 (63.4)
N2c	9 (11.8)	13 (16.9)	43 (20.9)	39 (19.0)
N3	2 (2.6)	3 (3.9)	7 (3.4)	5 (2.4)

Withdrawals

There were 13 withdrawals in the ND group and three in the surveillance group (*Table 10*).

Protocol deviations

Categories and frequencies of protocol deviations are given in *Table 11*. A number of patients were found to have been randomised on the basis of incorrect information. The tumour site category (oral, oropharyngeal, laryngeal, hypopharyngeal or occult) was inconsistent with the more detailed subsites. These errors were later corrected. Chemotherapy schedules were incorrectly recorded because of confusion over the three schedules containing cisplatin. Correction of these errors did not adversely affect the balance of the treatment allocation (see *Baseline characteristics*). The reasons why some patients in the ND arm did not undergo a ND are better described in *Neck dissections*. Twenty patients were randomised after they had already started CRT because the centre misunderstood the correct trial procedure.

Chemotherapy

Chemotherapy schedules planned at randomisation

Planned chemotherapy schedules are counted separately for patients in whom ND was planned to take place before CRT (*Table 12*) and for those in whom ND was planned to take place after CRT (*Tables 13* and *14*). Neoadjuvant schedules were more frequent when the planned timing was ND after CRT. A slightly higher proportion of patients in the ND arm whose ND was planned to be before CRT had a planned concomitant platin schedule.

Type of chemotherapy delivered

Chemotherapy details were received for 272 patients in the ND arm and 277 patients in the surveillance arm. Overall, 10 ND patients and five surveillance patients did not have chemotherapy for the reasons given in *Table 15*.

One patient in the ND arm, who was planned (at randomisation) to receive neoadjuvant and concomitant chemotherapy, received only concomitant chemotherapy. At the same time, two patients randomised to concomitant platinum received only neoadjuvant and concomitant therapy.

TABLE 10 Reasons for withdrawal from the trial

Reason	Trial arm (*n*)	
	Surveillance	ND
Ineligible	–	1
Patient did not want ND/wanted surveillance arm	–	4
Wished to consent to another clinical trial	–	1
Migrated to the USA for IMRT	–	1
Patient decision	1	2
Patient and clinician decision	1	–
Clinician decision	1	–
No reason given	–	4

IMRT, intensity-modulated radiotherapy.

TABLE 11 Protocol deviations

Description	Trial arm (*n*)		Total number
	Surveillance	ND	
Related to randomisation			
Randomised on the basis of incorrect ND timing	–	4	4
Randomised on the basis of incorrect T stage	1	–	1
Randomised on the basis of incorrect chemotherapy schedule	27	30	57
Randomised on the basis of incorrect tumour site	39	33	72
Subtotal	67	67	134
Others related to either trial arm			
Started CRT prior to randomisation (mostly one centre)	11	9	20
Timeline of response	–	1	1
Chest radiography not CT of chest	1	1	2
MRI standardised uptake value details	–	1	1
Different chemotherapy schedule to planned	2	–	2
Ineligible	2	2	4
Subtotal	16	14	30
Others related to surveillance arm only			
No PET–CT scan done	1	–	1
Watch and wait on partial response	1	–	1
Timeline of PET–CT	2	–	2
Partial response in nodes led to biopsy taken after PET–CT	2	–	2
Subtotal	6	–	6
Others related to ND arm only			
Randomised to ND arm but patient elected not to undergo ND	–	25	25
No surgery, clinician decision	–	9	9
ND surgery later than protocol timeline	–	2	2
Timing of surgery changed (before/after CRT)	–	1	1
Underwent PET–CT	–	1	1
Subtotal	–	38	38

TABLE 12 Planned chemotherapy schedules: randomised as planned ND before CRT

Schedule	Trial arm, *n* (%)	
	Surveillance	ND
Concomitant cisplatin or carboplatin	55 (72.4)	63 (81.8)
Concomitant cetuximab	5 (6.6)	5 (6.5)
Neoadjuvant and concomitant platinum	3 (3.9)	–
Neoadjuvant docetaxel, cisplatin and 5-fluorouracil and concomitant cisplatin	12 (15.8)	9 (11.7)
Other agreed (neoadjuvant carboplatin)	1 (1.3)	–
Total	76	77

TABLE 13 Planned chemotherapy schedules: randomised as planned ND after CRT

Schedule	Trial arm, *n* (%)	
	Surveillance	ND
Concomitant cisplatin or carboplatin	108 (52.4)	106 (51.7)
Concomitant cetuximab	9 (4.4)	9 (4.4)
Neoadjuvant and concomitant platinum	7 (3.4)	9 (4.4)
Neoadjuvant docetaxel, cisplatin and 5-fluorouracil and concomitant cisplatin	77 (37.4)	76 (37.1)
Other agreed (see *Table 14*)	5 (2.4)	5 (2.4)
Total	206	205

TABLE 14 Descriptions of other agreed chemotherapy

Chemotherapy received	Number of patients
Surveillance arm	
Neoadjuvant cisplatin and 5-fluorouracil, concomitant cisplatin	2
Cisplatin and 5-fluorouracil	1
TPF	2
ND arm	
Neoadjuvant docetaxel, cisplatin and 5-fluorouracil, concomitant cetuximab	1
Neoadjuvant cisplatin and 5-fluorouracil, concomitant cisplatin (2)	2
Induction cisplatin and 5-fluorouracil followed by concomitant cisplatin neoadjuvant and concomitant platinum	1
Neoadjuvant docetaxel and cisplatin and 5-fluorouracil, concomitant carboplatin	1

TABLE 15 Reasons why no chemotherapy was received

Reason for no chemotherapy	Trial arm (*n*)	
	Surveillance	ND
Refused	1	–
Patient unwell	–	2
Death	1	3
Did not want ND	–	2
Wanted surgery	1	–
Progression/metastases	1	1
Went to the USA for treatment	–	1
Withdrew	1	–
Not known	–	1

Individual drug details were available for 269 patients in the ND arm and 274 patients in the surveillance arm (*Table 16*). The type of chemotherapy delivered was very similar in the trial arms.

Chemotherapy dose information

Table 17 gives median doses received as reported on the CRT chemotherapy forms.

TABLE 16 Chemotherapy received

Chemotherapy received	Trial arm (*n*)	
	Surveillance	ND
Concomitant cisplatin or carboplatin schedule planned		
Concomitant (cisplatin or carboplatin)	154	153
Concomitant (cisplatin or carboplatin) and 5-fluorouracil	–	1
Concomitant (cisplatin or carboplatin) then changed to cetuximab	–	3
Concomitant cetuximab	1	2
TPF	2	–
Concomitant cetuximab schedule planned		
Concomitant cetuximab	12	13
Neoadjuvant and concomitant platin schedule planned		
Neoadjuvant (cisplatin or carboplatin) and 5-fluorouracil, concomitant (cisplatin or carboplatin)	8	7
Neoadjuvant (cisplatin or carboplatin) and 5-fluorouracil	1	2
Neoadjuvant (cisplatin or carboplatin)	1	–
TPF schedule planned		
TPF	71	61
Neoadjuvant (cisplatin or carboplatin), 5-fluorouracil and docetaxel	6	8
Neoadjuvant (cisplatin or carboplatin), 5-fluorouracil and docetaxel, concomitant cetuximab	8	6
Neoadjuvant (cisplatin or carboplatin)	–	1
Neoadjuvant docetaxel and concomitant cisplatin or carboplatin	–	1
Neoadjuvant (cisplatin or carboplatin), 5-fluorouracil and cetuximab, concomitant cetuximab	1	–
Neoadjuvant (cisplatin or carboplatin), cetuximab, 5-fluorouracil, docetaxel	–	1
Neoadjuvant (cisplatin or carboplatin) and 5-fluorouracil, concomitant (cisplatin or carboplatin)	3	4
Concomitant (cisplatin or carboplatin)	–	1
Other approved schedules planned		
Neoadjuvant (cisplatin or carboplatin) and 5-fluorouracil, concomitant (cisplatin or carboplatin)	4	3
TPF	–	2
Neoadjuvant (cisplatin or carboplatin), 5-fluorouracil and docetaxel, concomitant cetuximab	1	–
Neoadjuvant (cisplatin or carboplatin), 5-fluorouracil and docetaxel	1	–
Total	274	269

TABLE 17 Median chemotherapy doses

Schedule and drug	Trial arm			
	Surveillance		ND	
	n	Dose (mg), median (Q1, Q3)	*n*	Dose (mg), median (Q1, Q3)
Concomitant platin				
Cisplatin	136	340 (240, 420)	140	337 (240, 400)
Carboplatin (excluding AUC)[a]	19	780 (550, 1050)	22	645 (440, 960)
Schedule 2: concomitant cetuximab				
Concomitant cetuximab	11	3500 (2580, 4210)	13	3260 (1400, 3870)
Neoadjuvant and concomitant platin				
Concomitant cisplatin	7	300 (282, 320)	6	335 (120, 360)
Neoadjuvant cisplatin	8	236 (180, 266)	8	284 (210, 400)
Neoadjuvant 5-fluorouracil	8	7625 (4450, 12,450)	8	10500 (7260, 14,820)
Schedule 4: neoadjuvant TPF and concomitant platin				
Concomitant cisplatin	46	332 (183, 400)	47	312 (200, 400)
Concomitant carboplatin (when not AUC)	27	920 (740, 1310)	18	972 (650, 1270)
Cetuximab	9	2120 (1650, 2300)	6	2250 (2000, 2650)
Neoadjuvant cisplatin	84	373 (264, 450)	74	400 (300, 450)
Neoadjuvant docetaxel	81	390 (240, 440)	71	420 (300, 450)
Neoadjuvant 5-fluorouracil	86	14500 (6000, 18,000)	75	14600 (5700, 18,000)

AUC, area under the curve.
a Carboplatin doses recorded as AUC are omitted.

Radiotherapy

There were 270 completed case report forms for patients randomised to ND and 278 for patients randomised to PET–CT surveillance. Intensity-modulated radiotherapy (IMRT) was administered in 159 out of 263 (60%) ND patients and 173 out of 269 (64%) surveillance patients. Reasons why patients did not receive radiotherapy are given in *Table 18*.

There were few modifications to radiotherapy duration (*Table 19*). The duration of treatment in six ND patients and three surveillance patients was 5 or more days longer than planned.

The planned doses to primary in the two treatment arms were similar (*Table 20*).

Median and interquartile range (IQR) planned doses to involved neck nodes were also the same (54.0 Gy, IQR 50.0–54.0 Gy).

Neck dissections

Neck dissection was performed in 221 patients in the ND arm and in 54 patients in the surveillance arm.

TABLE 18 Reason for no radiotherapy

Reason	Trial arm (*n*)	
	Surveillance	ND
Refused (claustrophobic)	–	1
Patient unwell	–	1
Death	1	4
Did not want ND	–	2
Progression/metastases	2	2
Went to the USA for IMRT	–	1
Withdrew	1	–
Not known	–	1

TABLE 19 Reasons for modification of radiotherapy duration

Reason	Trial arm (*n*)	
	Surveillance	ND
Skin toxicity and mucositis	1	–
Myelosuppression	–	1
Pneumonia	–	1
Perforated duodenal ulcer	1	–
Did not attend	1	–
Weight loss: required replan	–	1
Neutropenia and infection	–	1
Admitted to hospital	–	1
Constipation/poor nutritional intake	–	1

TABLE 20 Planned radiotherapy schedules

Radiotherapy dose fractionation	Trial arm, *n* (%)	
	Surveillance	ND
68–70 Gy in 34 or 35 fractions	95 (34.7)	86 (32.2)
60–66 Gy in 30 fractions	144 (52.6)	144 (53.9)
55 Gy in 20 fractions	28 (10.2)	28 (10.5)
Other	7 (2.6)	9 (3.4)
Not known	4	3

Reproduced from the New England Journal of Medicine, Mehanna H, Wong W-L, McConkey CC, Rahman JK, Robinson M, Hartley AGJ, PET-CT Surveillance versus Neck Dissection in Advanced Head and Neck Cancer, 374, 1444–54. Copyright © (2016) Massachusetts Medical Society. Reprinted with permission.[62]

Patients in neck dissection arm who did not undergo a neck dissection

Of the 282 patients in the ND arm, 221 underwent a planned ND. In total, 25 refused to have surgery. A summary of reasons for the 61 who did not receive a planned ND is given in *Table 21*.

Summary of surveillance arm patients who had a neck dissection

According to the trial protocol, patients in the surveillance arm should undergo a ND if the PET–CT findings are positive or equivocal. Fifty-four patients underwent ND, with 48 (36 incomplete responders in nodes and 12 incomplete responders in primary and nodes) of them conforming to protocol in that PET–CT was positive or equivocal. Of the remaining six, two patients did not undergo PET–CT, in two patients PET–CT findings were equivocal or node positive in the central reviewer's PET–CT read and in two cases the reason is unknown.

Summary of surveillance arm patients who did not have a neck dissection

There were 228 patients in the surveillance arm who had no ND (*Table 22*).

Surgical details

Surgery forms were received for 275 patients who received ND surgery (*Table 23*).

From the above, it is important to note that, overall, the number of participants who experienced nerve sacrifice (which is morbid) was considerably lower in the surveillance arm (4 out of 220) than in the ND arm (22 out of 220); similar findings were obtained for sacrifice of the internal jugular vein and sternocleidomastoid muscle.

If the numbers of structures removed are summed (counting potentially one for left and one for right), the mean number of structures removed per ND is 0.59 in the ND arm and 0.56 in the surveillance arm. The mean number of structures removed in the ND before group is 0.78, whereas in the ND after group it is 0.49. Combining the five structures given, fewer patients whose planned ND was after CRT had one or more of the structures removed (see *Table 23*, $p = 0.02$). On examination of these data, this does not correlate with more advanced disease.

TABLE 21 Reasons for patients in ND arm not undergoing ND

Reason	ND planned (*n*)	
	Before CRT	After CRT
Clinical decision	–	3
Inoperable or unsuitable for surgery	1	1
Progression/metastases/residual disease	1	8
Patient unwell	1	6
Early death	-	4
Patient elected not to have ND	2	23
Went to the USA for IMRT	1	
Patient elected not to have ND after delay in finding operation list date	1	
Had PET–CT scan and CR	–	2
Patient wanted PET–CT arm	–	3
No information	1	3

TABLE 22 Reasons for patients in surveillance arm who had no ND

Reason	*n*
PET–CT, CRT, CR in primary and nodes	181
PET–CT, CRT, CR in nodes but not in primary	15
PET–CT, CRT, incomplete response in primary and nodes	7
Did not receive CRT	6
CR in primary, not in nodes	11
Retroperitoneal nodes	1
Nodes reported as CR, but after PET–CT reviewed reassigned as involved	1
New lung primary	1
Distant metastases	1
Neck node judged by MDT to be reactive	1
Recurred 5 weeks after post-CRT assessment	1
Metastases of liver and bone found at assessment	1
Biopsy confirmed no malignancy	1
Faint activity in lymph node on PET–CT but deemed to be a CR	1
MDT deemed nodal response as clear, CR	1
Biopsy, found only residuum of previously involved node	1
No information about response in primary	4
Had imaging other than PET–CT	2
Missed appointment and died	1
Had progressive disease	1
No response information having died early	2

Adapted from the New England Journal of Medicine, Mehanna H, Wong W-L, McConkey CC, Rahman JK, Robinson M, Hartley AGJ, PET-CT Surveillance versus Neck Dissection in Advanced Head and Neck Cancer, 374, 1444–54. Copyright © (2016) Massachusetts Medical Society. Reprinted with permission.[62]

Follow-up

Patients were due to be followed up by examination at the clinic at 6, 12 and 24 months after randomisation. Late requests for survival and recurrence information were made in the last year of the trial to obtain follow-up data for up to 5 years after randomisation.

Median follow-up for survival in the ND arm is 36.1 months (IQR 25.6–43.7 months) and for the surveillance arm is 35.6 months (IQR 29.2–42.5 months). The survival status at 24 months is unknown for 28 patients in the ND arm and 16 in the surveillance arm. Eleven patients in the ND arm and one in the surveillance arm have no follow-up, having, with one exception, withdrawn early. *Table 24* gives information on the timing of follow-up visits.

Serious adverse events

Summary

In the study, 282 SAEs were reported and are summarised in *Table 25*. There were 169 in the ND arm and 113 in the PET–CT arm; none was deemed unexpected by the chief investigator.

TABLE 23 Neck dissection details

ND details	ND, *n* (%)		Trial arm, *n* (%)	
	Before CRT (*N* = 69)	After CRT (*N* = 152)	ND (*N* = 221)	Surveillance (*N* = 54)
Laterality				
Unilateral left	26 (37.7)	62 (41.3)	88 (40.2)	23 (42.6)
Unilateral right	36 (52.2)	68 (45.3)	104 (47.5)	28 (51.9)
Bilateral	7 (10.1)	20 (13.3)	27 (12.3)	3 (5.6)
Not available	–	2	2	–
Type of ND				
Modified radical	39 (56.5)	57 (37.5)	96 (43.4)	19 (35.2)
Selective	29 (42.0)	87 (57.2)	116 (52.5)	30 (55.6)
Super selective	1 (1.4)	8 (5.3)	9 (4.1)	5 (9.3)
Type of incision				
J-shaped	15 (22.7)	39 (27.1)	54 (25.7)	14 (26.4)
Y-shaped (triradiate)	13 (19.7)	9 (6.3)	22 (10.5)	1 (1.9)
U-shaped (utility)	11 (16.7)	26 (18.1)	37 (17.6)	10 (18.9)
McFee	3 (4.5)	8 (5.6)	11 (5.2)	9 (17.0)
Other	24 (36.4)	62 (43.1)	86 (41.0)	19 (35.8)
Not available	3	8	11	1
Neck structures removed				
Internal jugular vein	20 (29.0)	24 (15.8)	44 (19.9)	14 (25.9)
Spinal accessory nerve	12 (17.4)	10 (6.6)	22 (10.0)	4 (7.4)
Sternocleidomastoid muscle	20 (29.0)	30 (19.7)	50 (22.6)	11 (20.4)
Common and internal carotid artery	2 (2.9)	2 (1.3)	4 (1.8)	–
Skin	3 (4.3)	1 (0.7)	4 (1.8)	1 (1.9)
Number of patients with any of the above structures removed	30 (43.5)	42 (27.6)	72 (32.6)	17 (31.5)

Adapted from the New England Journal of Medicine, Mehanna H, Wong W-L, McConkey CC, Rahman JK, Robinson M, Hartley AGJ, PET-CT Surveillance versus Neck Dissection in Advanced Head and Neck Cancer, 374, 1444–54. Copyright © (2016) Massachusetts Medical Society. Reprinted with permission.[62]

TABLE 24 Timing of follow-up visits by trial arm

Trial follow-up time point	Trial arm			
	Surveillance		ND	
	Number of patients	Time (months) from randomisation to examination median (IQR)	Number of patients	Time (months) from randomisation to examination median (IQR)
6 months	270	6.3 (5.9–7.1)	258	6.3 (5.8–7.3)
12 months	257	12.2 (11.7–12.9)	234	12.3 (11.8–13.1)
24 months	227	24.3 (23.7–25.3)	210	24.3 (23.8–25.3)

TABLE 25 Serious adverse events summary

	Trial arm, *n*			
	ND			Surveillance
	Planned			
SAE summary	Before CRT (*N* = 77)	After CRT (*N* = 205)	Total (*N* = 282)	Total (*N* = 282)
Number of SAEs	35	134	169[a]	113[a]
Number of patients with at least one SAE	25	87	112	89
Percentage of patients experiencing SAEs	32.5	42.4	39.7	31.6

a $p = 0.001$ for difference in number of SAEs between arms.

Tables 26 and *27* show that the most common SAE was hospitalisation or prolongation of hospitalisation due to CRT, which was very high for ND post CRT. The rates of other SAEs are similar in the ND before and after CRT trial arms and in the surveillance ND arm.

TABLE 26 Primary reasons for reporting SAEs

	Trial arm, *n*			
	ND			Surveillance
	Planned			
Primary reasons for reporting (> 1 allowed per patient)	Before CRT (*N* = 77)	After CRT (*N* = 205)	Total (*N* = 282)	Total (*N* = 282)
Death	2	1	3	3
Life-threatening event	2	13	15	11
Inpatient hospitalisation or prolongation of existing hospitalisation	29	129	158	104
Persistent or significant disability/incapacity	3	6	9	10
Other medically significant reason for reporting	2	4	6	5

TABLE 27 Causality of serious events

	Trial arm, *n*			
	ND			Surveillance
	Planned			
Causality	Before CRT	After CRT	Total	Total
Surgery	4	6	10	5
CRT	22	88	110	75
CRT and surgery	–	2	2	–
CRT and other	–	4	4	1
Other	9	34	43	32

Of the SAEs with hospitalisation as a result of to CRT in the ND after CRT group, 67 out of 71 patients (94.4%) had their SAE before surgery (*Table 28*).

The highest symptom grades of the events are given comparing trial arms in *Table 29* and comparing ND before CRT with ND after CRT in *Table 30*. No difference is seen between groups in either case ($p = 0.99$ and $p = 0.42$ for trend).

TABLE 28 Outcome of SAEs

	Trial arm, *n*			
	ND			Surveillance
	Planned			
Outcomes	Before CRT (*N* = 35 events)	After CRT (*N* = 134 events)	Total	Total
Resolved	27	107	134	94
Not yet resolved	6	23	29	15
Deaths	2	4	6	4

TABLE 29 Symptom grades of SAEs by trial arm

	Trial arm, number of SAEs (%)	
Highest symptom grade of event	Surveillance (*N* = 89)	ND (*N* = 112)
5	1 (0.4)	1 (0.4)
4	8 (2.8)	15 (5.3)
3	64 (22.7)	71 (25.2)
2	10 (3.5)	18 (6.4)
1	6 (2.1)	7 (2.5)

No difference is seen in the grades of the events ($p = 0.99$ for trend).

TABLE 30 Symptom grades of SAEs: ND before and after CRT

	ND, number of SAEs (%)	
Highest symptom grade of event	Before CRT (*N* = 77)	After CRT (*N* = 205)
5	– (–)	1 (0.5)
4	4 (5.2)	11 (5.4)
3	13 (16.9)	58 (28.3)
2	6 (7.8)	12 (5.9)
1	2 (2.6)	5 (2.4)
0, that is, no SAE	52 (67.5)	118 (57.6)

$p = 0.09$ for trend; $p = 0.42$ excluding last line.

Overall survival

Median follow-up for survival was 35.6 months (IQR 29.2–42.5 months) in the surveillance arm and 36.1 months (IQR 25.6–43.7 months) in the ND arm.

There have been 123 deaths reported in the trial: 62 in the ND arm and 61 in the surveillance arm. Of these, 92 occurred within 2 years of randomisation. One death in the surveillance arm occurred > 5 years after randomisation and is not included in the time-to-event analysis.

Overall survival is plotted by treatment arm (*Figure 4*). The 2-year OS rate for the trial was 83.2% (95% CI 80.1% to 86.4%). This was higher than that expected at the start of the trial (75%). Consequently, the retrospective power of the study with 122 deaths was 72%.

The 2-year OS in the ND arm was 81.5% (95% CI 76.9% to 86.2%) and in the PET–CT surveillance arm was 84.9% (95% CI 80.7% to 89.1%).

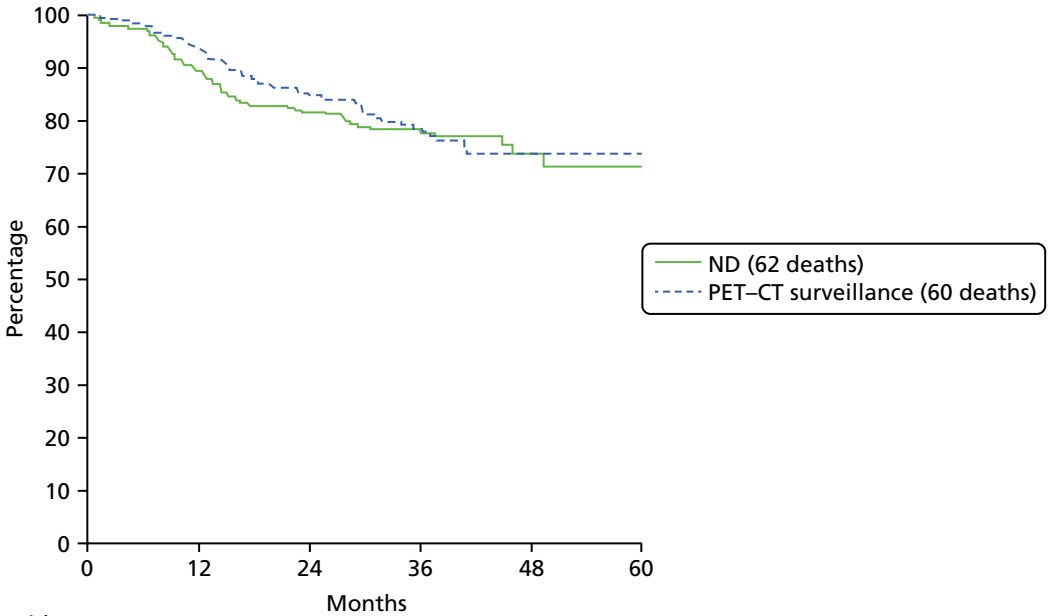
Primary outcome overall survival

The prespecified analysis takes account of timing of planned ND by stratification (planned ND before CRT and planned ND after CRT). These are plotted separately in *Figures 5* and *6*, respectively.

Analysing by proportional hazards model (*Table 31*) stratified by intended timing of ND before or after CRT, the HR is 0.924 (95% CI 0.648 to 1.318) in favour of the surveillance arm. The limit for non-inferiority in the trial design was set at a HR of 1.50, which lies at the 99.63th percentile of the estimated CI for the HR, which can be interpreted as a one-sided *p*-value of 0.0037, that is, indicating non-inferiority (HR < 1.5). The conventional two-sided *p*-value for difference between treatments is 0.6621.

Further analyses of overall survival

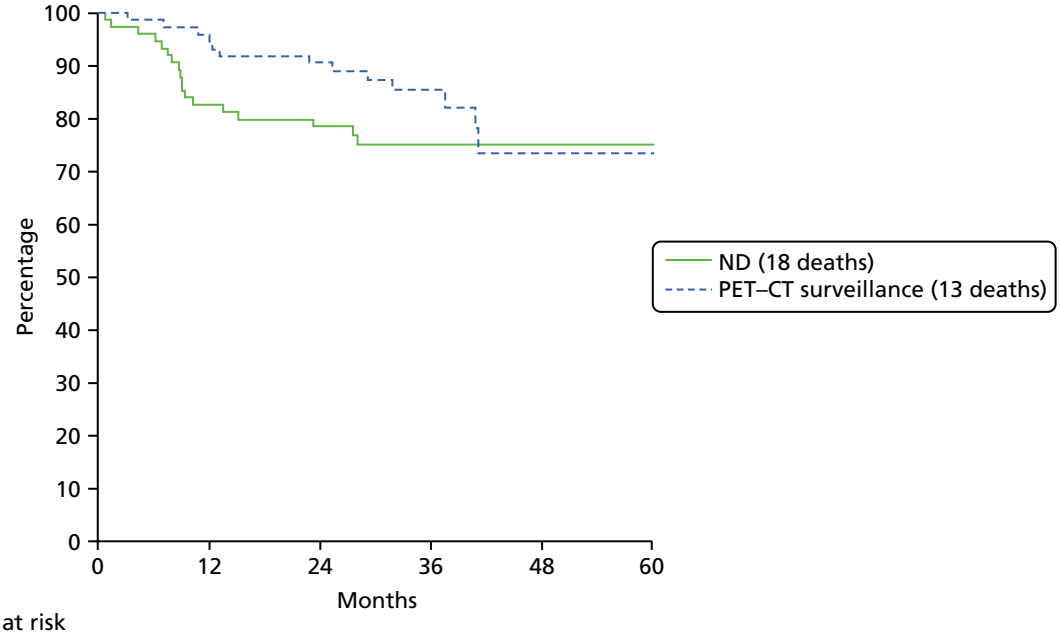
The early drop in the Kaplan–Meier curve of the group that had a planned ND before CRT suggests some non-proportionality of hazards between treatment groups. This is confirmed by testing an interaction between



Number at risk

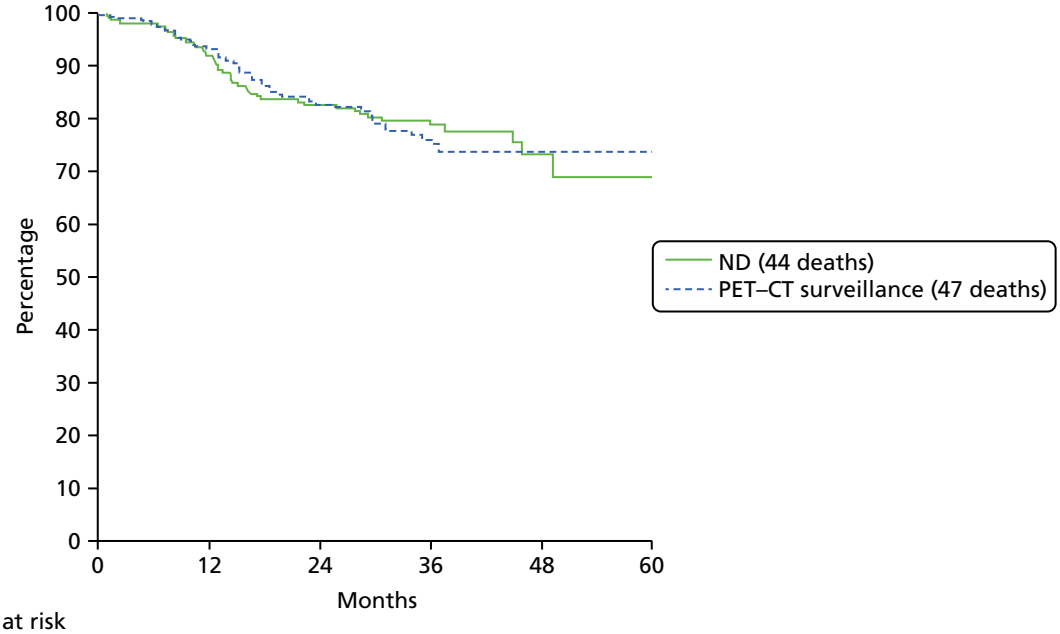
ND	282	243	204	118	32	8
PET–CT surveillance	282	259	224	110	33	6

FIGURE 4 Overall survival by treatment arm. Reproduced from the New England Journal of Medicine, Mehanna H, Wong W-L, McConkey CC, Rahman JK, Robinson M, Hartley AGJ, PET-CT Surveillance versus Neck Dissection in Advanced Head and Neck Cancer, 374, 1444–54. Copyright © (2016) Massachusetts Medical Society. Reprinted with permission.[62]



Number at risk						
ND	77	62	51	31	12	6
PET–CT surveillance	76	70	65	32	10	4

FIGURE 5 Overall survival by treatment arm for planned ND before CRT. Reproduced from the New England Journal of Medicine, Mehanna H, Wong W-L, McConkey CC, Rahman JK, Robinson M, Hartley AGJ, PET-CT Surveillance versus Neck Dissection in Advanced Head and Neck Cancer, 374, 1444–54. Copyright © (2016) Massachusetts Medical Society. Reprinted with permission.[62]



Number at risk						
ND	205	181	153	87	20	2
PET–CT surveillance	206	189	159	78	23	2

FIGURE 6 Overall survival by treatment arm for planned ND after CRT. Reproduced from the New England Journal of Medicine, Mehanna H, Wong W-L, McConkey CC, Rahman JK, Robinson M, Hartley AGJ, PET-CT Surveillance versus Neck Dissection in Advanced Head and Neck Cancer, 374, 1444–54. Copyright © (2016) Massachusetts Medical Society. Reprinted with permission.[62]

TABLE 31 Primary outcome: OS

OS proportional hazards model	Number of events	Treatment HR estimate (95% CI)	HR 1.50 (%)
Stratified by ND pre/post (prespecified analysis of primary outcome)	122	0.924 (0.648 to 1.318)	99.63

treatment and log-(time) (p-value for the time interaction is 0.03). No evidence of non-proportionality is seen in the planned ND after CRT stratum ($p = 0.38$). An alternative analysis to the Cox's model is stratifying analysis into two risk periods, namely the first year and rest of the period at risk. The result of this additional unplanned analysis gives a very similar HR of 0.922 (instead of 0.924 in *Table 31*).

Adjusting for centre, T stage, N stage, site of primary, chemotherapy schedule, sex and age gives a very similar result, as does further adjustment for p16 status on a reduced sample (*Table 32*).

p16 status was highly prognostic, but there was no difference between p16-positive and p16-negative status in terms of the trial result, as shown in *Table 32* and *Figures 7* and *8*.

The HRs in *Figure 9* are stratified by planned ND timing (before or after CRT). Some of the groups are small and differences between them should be interpreted cautiously. There is some evidence that the ND arm did better in the T stage 1/2 group than in the T3–T4 group. The small group of females in the trial did particularly well in the surveillance arm (treatment interaction $p = 0.01$), although with a total of only 18 deaths this is not a reliable result. There was no difference in the trial result by p16 status, which was highly prognostic. In unplanned analyses, the ND arm did worse in the first year at risk, and no difference was seen in the trial result by radiotherapy dose, IMRT versus conformal radiotherapy or high- versus low-recruiting centres.

Mortality by cause

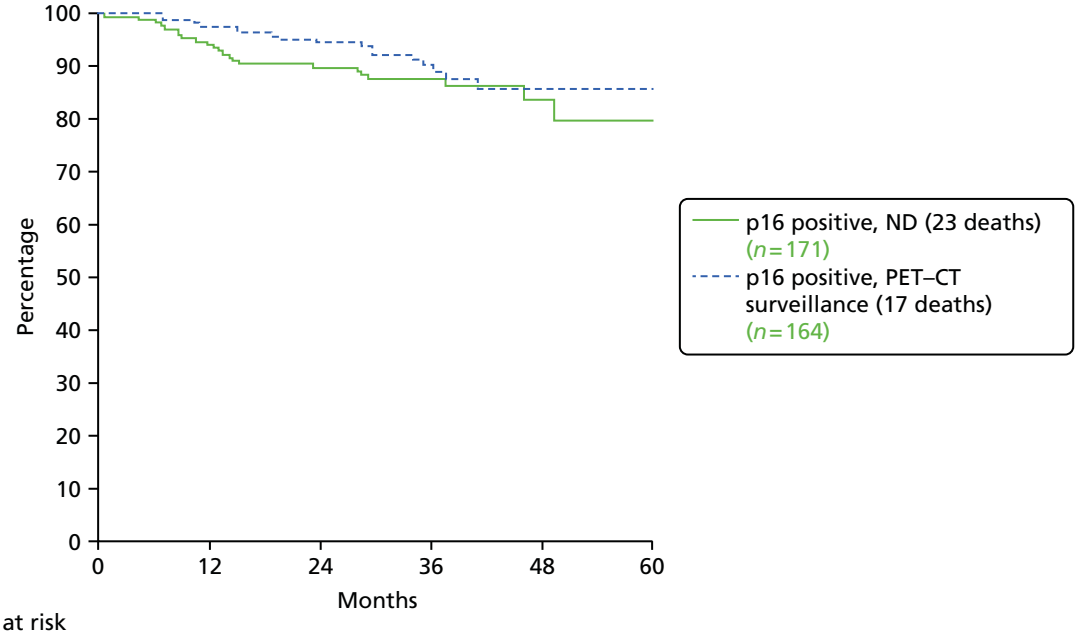
Deaths attributable to head and neck cancer

The causes of death are summarised in *Table 33*. These include complications attributable to surgery or CRT and persistent, recurrent or metastatic H&N cancer at death. There were 92 deaths due to H&N cancer and 29 deaths from other causes (detailed in *Table 34*).

TABLE 32 Overall survival: Cox's models adjusted and subgroups

OS proportional hazards model	Number of events	Treatment HR estimate (95% CI)	HR 1.50 (%)
Stratified by ND pre/post adjusted for centre, T stage, N stage, tumour site, chemotherapy schedule, sex and age	122	0.835 (0.571 to 1.222)	99.87
Stratified by ND pre/post, adjusted as above plus p16 status (categorised as positive/negative/not known)	122	0.759 (0.513 to 1.124)	99.97
ND before CRT subgroup	31	0.664 (0.325 to 1.356)	98.7
ND after CRT subgroup	91	1.033 (0.685 to 1.558)	96.2
p16-positive subgroup stratified by ND pre/post	40	0.736 (0.393 to 1.378)	98.7
p16-negative subgroup stratified by ND pre/post	58	0.982 (0.583 to 1.654)	94.4

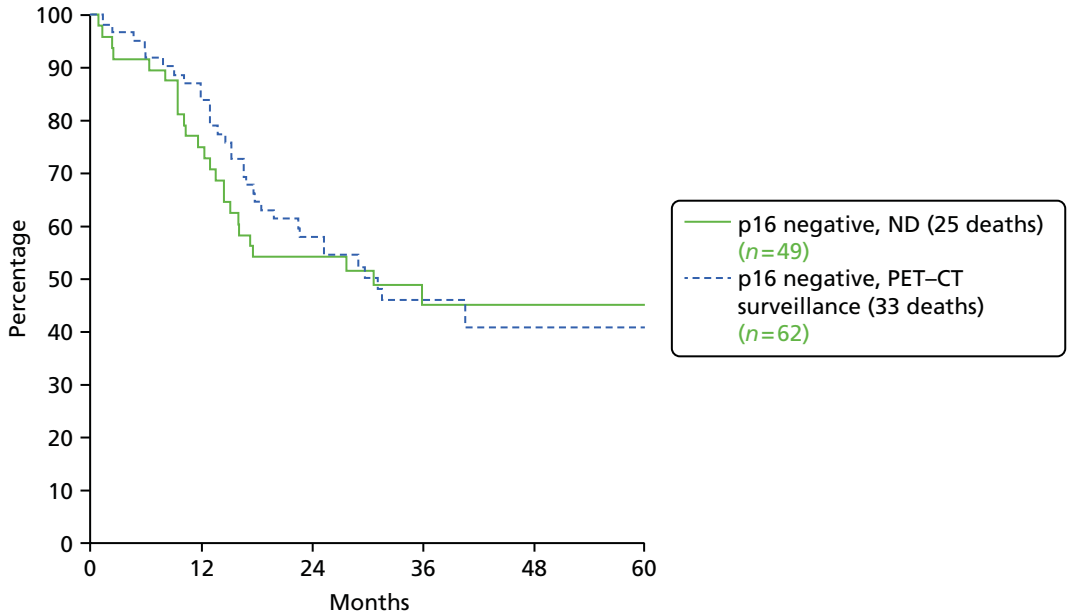
Adapted from the New England Journal of Medicine, Mehanna H, Wong W-L, McConkey CC, Rahman JK, Robinson M, Hartley AGJ, PET-CT Surveillance versus Neck Dissection in Advanced Head and Neck Cancer, 374, 1444–54. Copyright © (2016) Massachusetts Medical Society. Reprinted with permission.[62]



Number at risk

ND	171	157	140	81	23	5
PET–CT surveillance	164	157	145	78	24	4

FIGURE 7 Overall survival: p16 positive.



Number at risk

ND	49	36	22	14	3	2
PET–CT surveillance	62	52	34	17	5	1

FIGURE 8 Overall survival: p16 negative.

Head and neck cancer mortality: cumulative incidence

There is little difference between the trial arms in the incidence of H&N cancer-related deaths (*Figure 10*).

The 2-year cumulative incidence rates for H&N mortality are 13.66% (95% CI 9.88% to 18.04%) in the ND arm and 12.25% (95% CI 8.72% to 16.42%) in the surveillance arm. A Gray's test for difference between cumulative incidence functions gives a *p*-value of 0.7992.

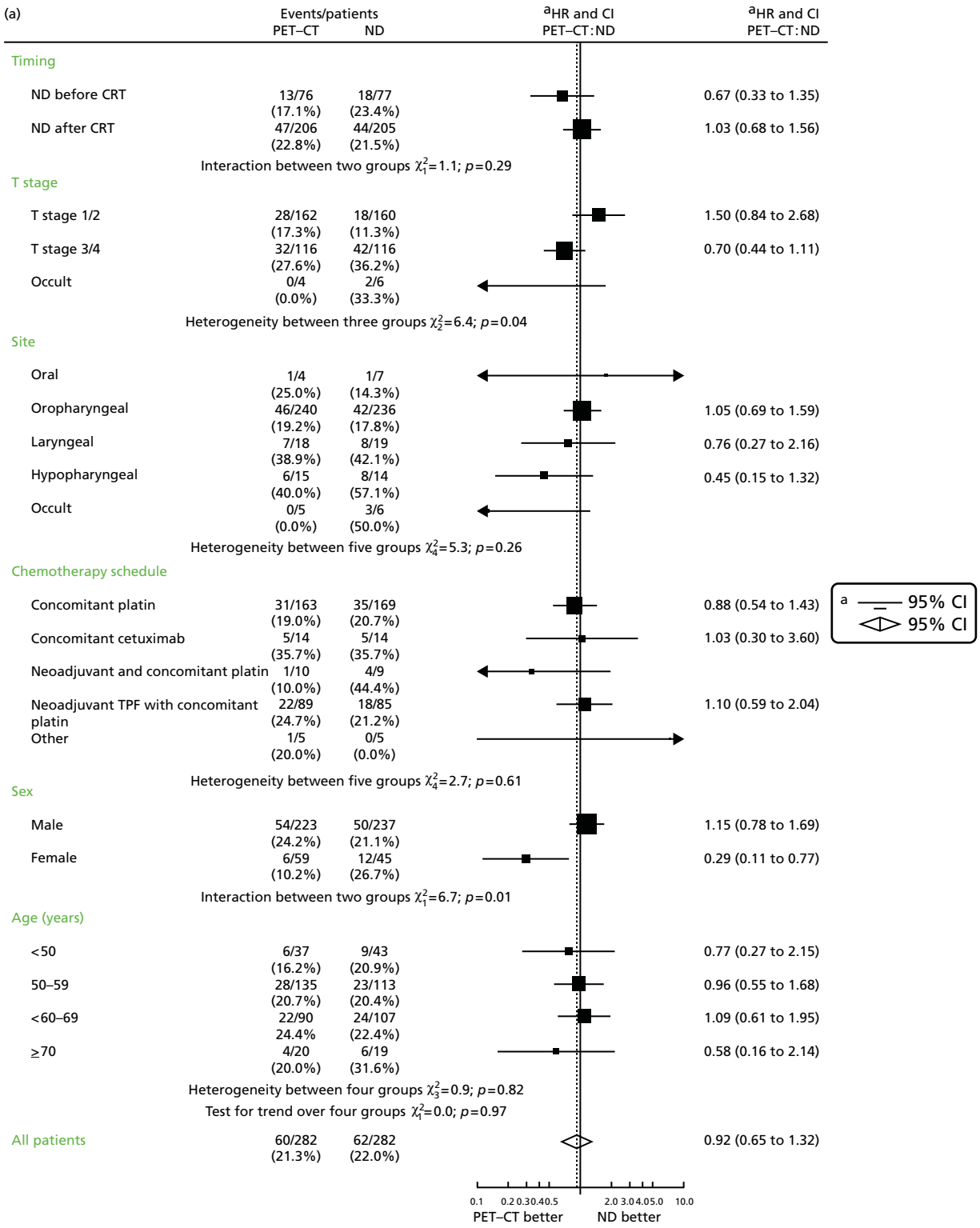


FIGURE 9 Overall survival of subgroups. Method as described in Early Breast Cancer Trialists' Collaborative Group. *Treatment of Early Breast Cancer*. Volume 1. *Worldwide Evidence 1985–1990*. Oxford: Oxford University Press; 1990.[57] (a) Overall survival of subgroups 1; and (b) overall survival of subgroups 2. Adapted from the New England Journal of Medicine, Mehanna H, Wong W-L, McConkey CC, Rahman JK, Robinson M, Hartley AGJ, PET-CT Surveillance versus Neck Dissection in Advanced Head and Neck Cancer, 374, 1444–54. Copyright © (2016) Massachusetts Medical Society. Reprinted with permission.[62] (*continued*)

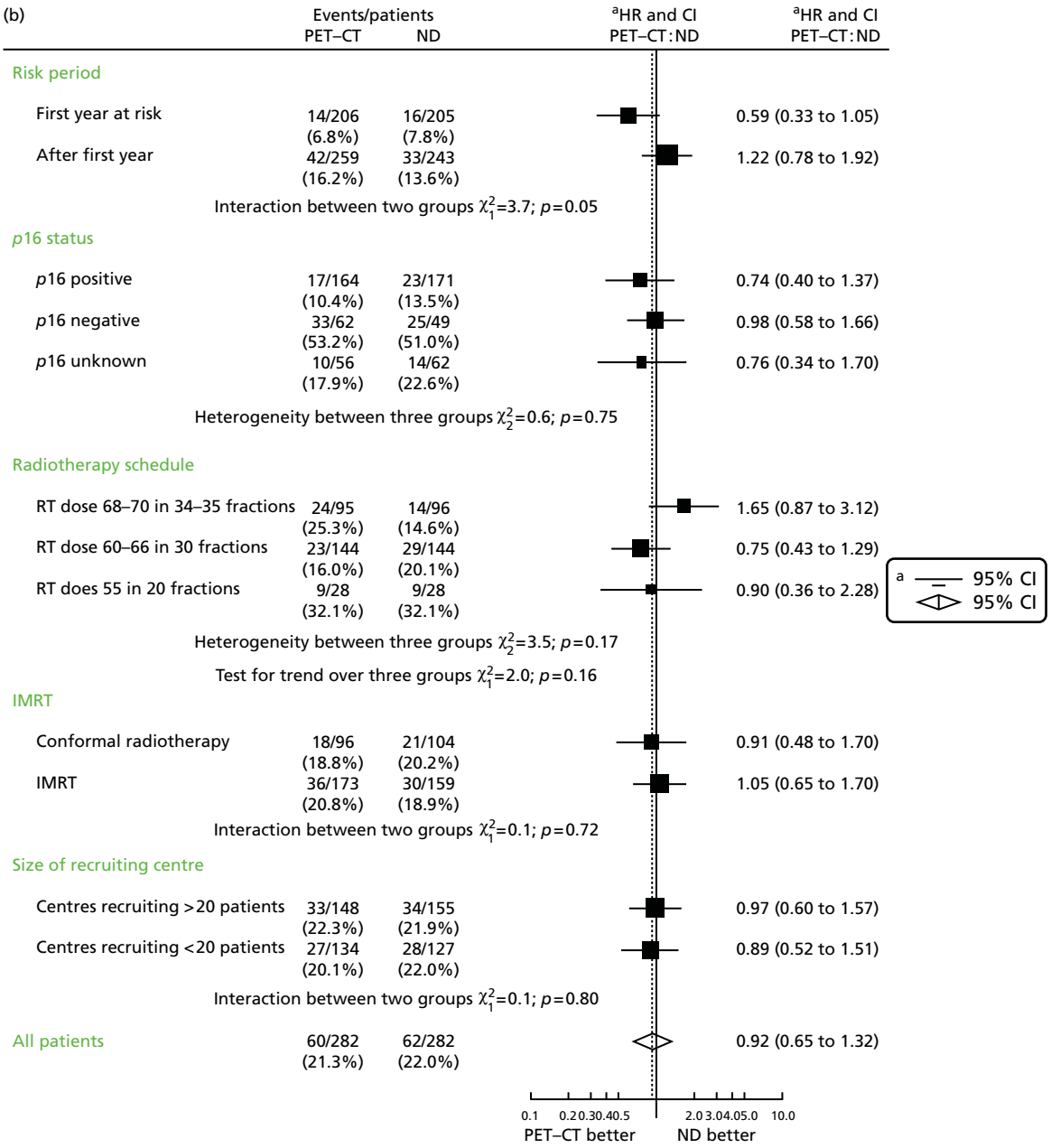


FIGURE 9 Overall survival of subgroups. Method as described in Early Breast Cancer Trialists' Collaborative Group. *Treatment of Early Breast Cancer*. Volume 1. *Worldwide Evidence 1985–1990*. Oxford: Oxford University Press; 1990.[57] (a) Overall survival of subgroups 1; and (b) overall survival of subgroups 2. Adapted from the New England Journal of Medicine, Mehanna H, Wong W-L, McConkey CC, Rahman JK, Robinson M, Hartley AGJ, PET-CT Surveillance versus Neck Dissection in Advanced Head and Neck Cancer, 374, 1444–54. Copyright © (2016) Massachusetts Medical Society. Reprinted with permission.[62]

Other cause mortality: cumulative incidence

The difference in non-cancer deaths is also small (*Figure 11*).

The 2-year cumulative incidence rates for other causes of mortality are 4.80% (95% CI 2.68% to 7.82%) in the ND arm and 2.88% (95% CI 1.36% to 5.37%) in the surveillance arm. A Gray's test for difference between cumulative incidence functions gives a *p*-value of 0.4124.

TABLE 33 Causes of death

	Randomised as ND policy (*n*)					
	Before CRT		After CRT		Total (*n*)	
Cause of death	Surveillance	ND	Surveillance	ND	Surveillance	ND
H&N cancer	11	13	37	31	48	44
Other causes	3	5	10	11	13	16
Not known	–	–	–	2	–	2
Total	14	18	47	44	61	62

TABLE 34 Description of non-H&N cancer causes

Arm	Cause
PET–CT	
Before CRT stratum	Cardiac failure
	Malignant plasma cell neoplasm: extramedullary plasmacytoma
	Pneumonia
After CRT stratum	Road traffic accident
	New primary non-small cell lung cancer
	Extensive bilateral consolidation secondary to severe infection
	Second primary lung
	Bronchopneumonia: squamous cell carcinoma oropharynx
	Severe depression
	Post-operative complications for second primary oesophagus
	Chest sepsis
	Unknown primary: probably lung, widespread metastases
	Renal cancer, cardiovascular and respiratory failure. Low anterior resection plus ileostomy
ND	
Before CRT	Pneumonia
	Aspiration pneumonia
	Glioblastoma
	Bilateral pneumonia
	Infection
After CRT	Unknown
	Patient on dialysis, died of renal failure
	Metastatic non-small cell lung cancer
	Left ventricular failure, myocardial infarction
	Second primary
	Sudden death at home, no recurrence of H&N cancer
	Neutropenic sepsis as a result of colitis
	Bronchopneumonia
	Cerebrovascular accident
	Coronary artery atherosclerosis
	Not known as patient moved away
	Aspiration pneumonia

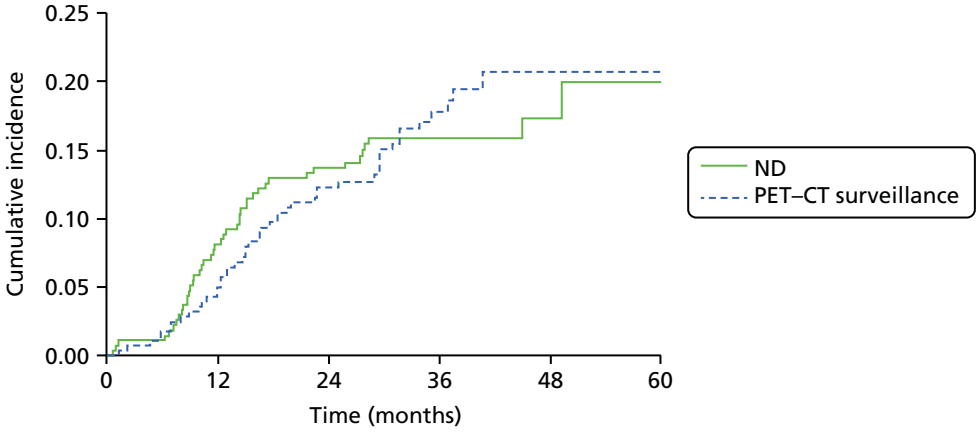


FIGURE 10 Mortality attributable to H&N cancer.

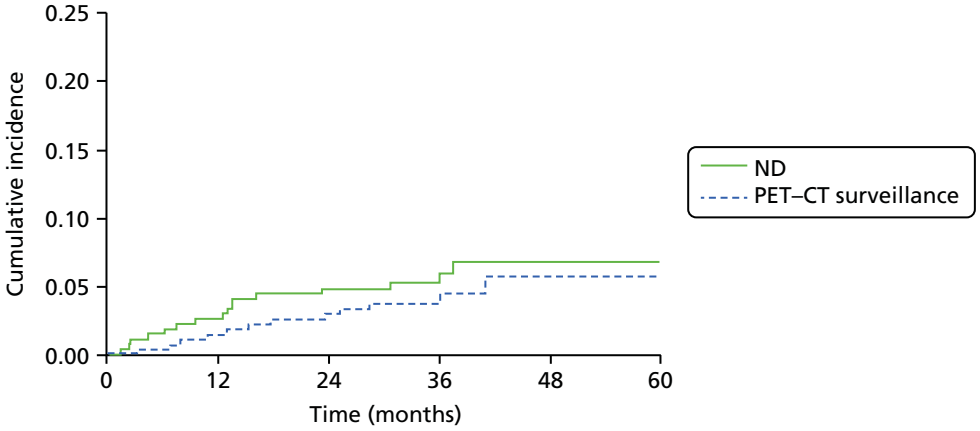


FIGURE 11 Mortality attributable to causes other than H&N cancer.

Recurrences

Information on recurrences and persistent disease is obtained mainly from recurrence forms, and also via cause of death information. Recurrence is defined using the date of completion of radiotherapy. Disease that is apparent < 3 months after radiotherapy is defined as persistent disease. Disease that is apparent later than 3 months is defined as a recurrence. Frequencies of persistent disease and reported sites of recurrence are shown in *Table 35*.

For Kaplan–Meier analysis of recurrence (*Figure 12*), the time of recurrence is taken as the time from completion of radiotherapy to reported recurrence or to cancer death if there is no reported recurrence. Consequently, the time at risk for some patients starts after the ND, while for others it is before ND and others still have no ND. The 2-year recurrence rates are 16.0% (95% CI 11.4% to 20.6%) in the ND arm and 14.4% (95% CI 10.1% to 18.7%) in the surveillance arm.

Four of the six surveillance patients with recurrence in neck nodes were complete responders on the post-CRT PET–CT scan.

Two of the three surveillance patients with recurrence in the neck had a CR only on post-CRT PET–CT.

TABLE 35 Recurrences and persistent disease

Recurrence/persistent disease	Trial arm, *n*	
	Surveillance	ND
Total reports of recurrence/persistent disease	59	52
Reports via recurrence forms	54	46
Reports via cause of death only	5	6
Persistent disease (< 3 months after radiotherapy completion)	14	10
Recurrence (> 3 months after radiotherapy completion)	45	42
Recurrence in primary	21	16
In primary only	15	12
Recurrence in neck nodes	6	2
In neck nodes only	3	1
Recurrence site distant	21	23
Distant only	17	18
Site unknown	4	6

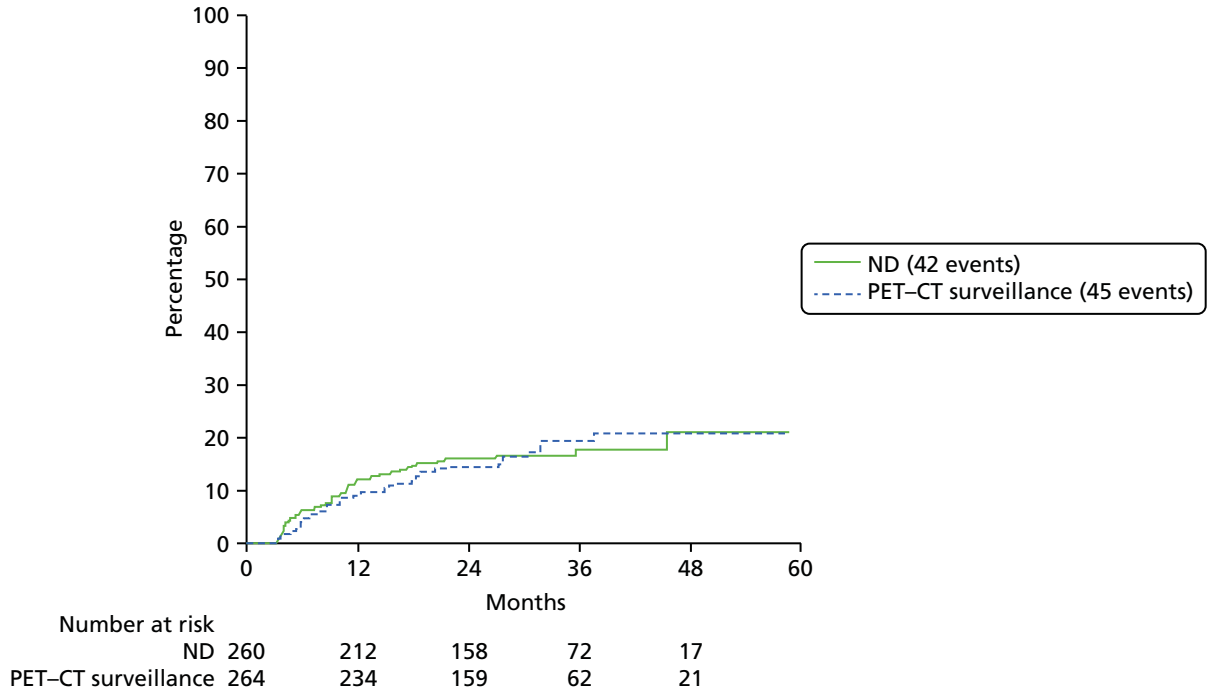


FIGURE 12 Time to recurrence by treatment arm.

Complications of neck dissection

Information on surgical complications was supplied for 220 of the 221 NDs in the ND arm and 53 of 54 in the surveillance arm. *Table 36* separates the results by ND pre/post CRT planned at the time of randomisation. The overall complication rate per ND is 0.38 in the ND arm and 0.42 in the surveillance arm.

TABLE 36 Complications of ND surgery by trial arm

Surgical complication	Trial arm, *n* (rate) [95% CI]			
	ND (*N* = 220)		Surveillance (*N* = 53)	
	Pre CRT (*n* = 69)	Post CRT (*n* = 151)	Pre CRT (*n* = 16)	Post CRT (*n* = 37)
No complications	50 (72.46%) [60.89% to 81.67%]	118 (78.15%) [70.86% to 84.03%]	10 (62.50%) [38.53% to 81.63%]	24 (64.86%) [48.70% to 78.23%]
General				
Chest infection	1 (1.45%) [0.00% to 8.52%]	3 (1.99%) [0.42% to 5.94%]	–	–
Urinary tract infection	–	–	–	2 (5.41%) [0.57% to 18.63%]
Septicaemia	–	1 (0.66%) [0.00% to 4.03%]	–	–
Other general (see *Table 37*)	–	2 (1.32%) [0.06% to 5.01%]	–	1 (2.70%) [0.00% to 15.05%]
Local				
Breach of tumour	1 (1.45%) [0.00% to 8.52%]	–	–	–
Intraoperative chyle leak	1 (1.45%) [0.00% to 8.52%]	1 (0.66%) [0.00% to 4.03%]	–	2 (5.41%) [0.57% to 18.63%]
Wound infection	3 (4.35%) [0.99% to 12.52%]	9 (5.96%) [3.02% to 11.09%]	1 (6.25%) [0.00% to 30.31%]	3 (8.11%) [2.06% to 22.03%]
Postoperative chyle leak	1 (1.45%) [0.00% to 8.52%]	3 (1.99%) [0.42% to 5.94%]	1 (6.25%) [0.00% to 30.31%]	–
Vagal palsy	1 (1.45%) [0.00% to 8.52%]	–	–	–
Marginal mandibular nerve palsy	5 (7.25%) [2.77% to 16.23%]	3 (1.99%) [0.42% to 5.94%]	1 (6.25%) [0.00% to 30.31%]	–
Postoperative haematoma	3 (4.35%) [0.99% to 12.52%]	5 (3.31%) [1.21% to 7.72%]	1 (6.25%) [0.00% to 30.31%]	–
Breakdown of wound	2 (2.90%) [0.20% to 10.57%]	6 (3.97%) [1.65% to 8.59%]	–	1 (2.70%) [0.00% to 15.05%]
Shoulder movement disability	4 (5.80%) [1.85% to 14.40%]	10 (5.96%) [3.02% to 11.09%]	–	3 (8.11%) [2.06% to 22.03%]
Hypoglossal nerve palsy	–	2 (1.32%) [0.06% to 5.01%]	1 (6.25%) [0.00% to 30.31%]	2 (5.41%) [0.57% to 18.63%]
Seroma	1 (1.45%) [0.00% to 8.52%]	5 (3.31%) [1.21% to 7.72%]	–	1 (2.70%) [0.00% to 15.05%]
Other local (see *Table 37*)	1 (1.45%) [0.00% to 8.52%]	9 (5.96%) [3.02% to 11.09%]	1 (6.25%) [0.00% to 30.31%]	1 (2.70%) [0.00% to 15.05%]
Total number of complications	24	59	6	16

The complications given in the table as 'other local' and 'other general' are as in *Table 37*.

In *Table 38*, complication counts are given by whether or not the NDs were actually performed before or after CRT.

Quality of life

Completion of quality-of-life questionnaires

The total return rate of quality-of-life questionnaires, based on the number that could have been received if all surviving patients completed all of them, was 76.4% in the ND arm and 80.6% in the surveillance arm. Generally, the return rates (*Table 39*) were at least 70% at each time point and they were returned at the required times.

The patients who responded to the questionnaire at 24 months were broadly similar in terms of baseline characteristics to those living patients from whom there was no response, but the non-responders had slightly worse survival beyond 24 months.

Questionnaires in both trial arms were completed at the designated times (*Table 40*). The questions were well answered, allowing most scales to be computable in the majority of cases (*Table 41*). The questionnaires were well completed, with most scales computable in the majority of cases.

Baseline quality-of-life questionnaire scores

The quality-of-life, H&N function and dysphagia-related scores were well balanced between the treatment arms (*Table 42*).

TABLE 37 Complications of ND surgery, other

Arm	ND pre/post	Other general complication
ND	Pre	Fistula
ND	Post	Throat pain
ND	Post	Tracheostomy
ND	Post	Other nerve damage
ND	Post	Reduced oral intake
ND	Post	Oedema
ND	Post	Other nerve damage
ND	Post	Reduced sensation over right ear
ND	Post	Oedema
ND	Post	Infection
ND	Post	Dysphagia
ND	Post	Oedema
PET–CT	Pre	Fistula
PET–CT	Post	Intraoperative possible air embolism unconfirmed
PET–CT	Post	Required continuous pressure airway support for 1 day

TABLE 38 Complications of ND surgery as received that are not intention to treat

Surgical complication	Trial arm, *n* (rate) [95% CI]		
	ND arm		Surveillance arm
	ND performed pre CRT (*N* = 64)	ND performed post CRT (*N* = 152)	ND performed (*N* = 54)
No complications	49 (76.56%) [64.76% to 85.35%]	118 (77.63%) [70.34% to 83.56%]	34 (62.96%) [49.60% to 74.60%]
General			
Chest infection	1 (1.56%) [0.00% to 9.14%]	3 (1.97%) [0.41% to 5.90%]	–
Urinary tract infection	–	–	2 (3.70%) [0.30% to 13.26%]
Septicaemia	–	1 (0.66%) [0.00% to 4.00%]	–
Other general	–	2 (1.32%) [0.06% to 4.97%]	1 (1.85%) [0.00% to 10.69%]
Local			
Intraoperative chyle leak	1 (1.56%) [0.00% to 9.14%]	1 (0.66%) [0.00% to 4.00%]	2 (3.70%) [0.30% to 13.26%]
Wound infection	1 (1.56%) [0.00% to 9.14%]	9 (5.92%) [3.00% to 11.02%]	4 (7.41%) [2.42% to 18.05%]
Postoperative chyle leak	1 (1.56%) [0.00% to 9.14%]	3 (1.97%) [0.41% to 5.90%]	1 (1.85%) [0.00% to 10.69%]
Vagal palsy	1 (1.56%) [0.00% to 9.14%]	–	–
Marginal mandibular nerve palsy	5 (7.81) [3.00% to 17.4%]	3 (1.97%) [0.41% to 5.90%]	1 (1.85%) [0.00% to 10.69%]
Postoperative haematoma	2 (3.13%) [0.23% to 11.33%]	5 (3.29%) [1.21% to 7.68%]	1 (1.85%) [0.00% to 10.69%]
Breakdown of wound	1 (1.56%) [0.00% to 9.14%]	6 (3.95%) [1.63% to 8.53%]	1 (1.85%) [0.00% to 10.69%]
Shoulder movement disability	4 (6.25%) [2.01% to 15.44%]	10 (6.58%) [3.48% to 11.82%]	3 (5.56%) [1.32% to 15.70%]
Hypoglossal nerve palsy	–	2 (1.32%) [0.06% to 4.97%]	3 (5.56%) [1.32% to 15.70%]
Seroma	1 (1.56%) [0.00% to 9.14%]	5 (3.29%) [1.21% to 7.68%]	1 (1.85%) [0.00% to 10.69%]
Other local	1 (1.56%) [0.00% to 9.14%]	9 (5.92%) [3.00% to 11.02%]	2 (3.70%) [0.30% to 13.26%]
Total number of complications	19	59	22

TABLE 39 Questionnaire received

	Trial arm	
Quality-of-life questionnaire status	Surveillance	ND
Total randomised	282	282
Number of quality-of-life questionnaires received	1	5
Baseline questionnaire		
Received (% of total alive)	262 (92.9)	261 (92.6)
Not received	19	16
2 weeks after CRT		
Received (% of total alive)	211 (75.4)	196 (70.8)
Not received	67	71
Died	2	5
Withdrawn from trial	1	5
6 months after randomisation		
Received (% of total alive)	208 (76.5)	183 (68.3)
Not received	61	71
Died	10	14
Withdrawn from trial	2	9
12 months after randomisation		
Received (% of total alive)	197 (76.1)	181 (72.7)
Not received	58	53
Died	23	33
Withdrawn from trial	3	10
24 months after randomisation		
Received (% of total alive)	195 (81.9)	179 (76.8)
Not received yet	38	39
Died	44	49
Withdrawn from trial	4	10

TABLE 40 Timing of questionnaires

	Time from randomisation to completion, median (IQR)	
Time point	Surveillance	ND
2 weeks after CRT	87 days (69–117 days)	95 days (72–117 days)
6 months	6.2 months (5.8–7.1 months)	6.4 months (5.8–7.3 months)
12 months	12.2 months (11.9–13.1 months)	12.4 months (11.9–13.2 months)
24 months	24.5 months (23.9–25.8 months)	24.5 months (23.9–25.7 months)

TABLE 41 Completeness of quality-of-life questionnaires

Completeness	Trial arm, *n* (%)		Total, *N* (%)
	Surveillance	ND	
EORTC QLQ-C30			
All scales computable	1000 (93.2)	928 (92.8)	1928 (93.0)
One scale (out of 15) missing	54 (5.0)	52 (5.2)	106 (5.1)
More than one scale missing	19 (1.8)	20 (2.0)	39 (1.9)
EORTC QLQ-H&N35			
All scales computable	865 (80.5)	818 (81.9)	1683 (81.2)
One scale (out of 18) missing	142 (13.2)	115 (11.5)	257 (12.4)
More than one scale missing	67 (6.2)	66 (6.6)	133 (6.4)
MDADI dysphagia questionnaire			
All scales computable	1023 (95.3)	949 (94.9)	1972 (95.1)
One scale (out of five) missing	19 (1.8)	18 (1.8)	37 (1.8)
More than one scale missing	31 (2.9)	33 (3.3)	64 (3.1)
EQ-5D (3L)			
All items present	1008 (93.9)	937 (93.7)	1945 (93.8)
At least one item missing	65 (6.1)	63 (6.3)	128 (6.2)

3L, three levels.

TABLE 42 Baseline quality-of-life scores

Quality-of-life scale	Trial arm, mean (SD)	
	Surveillance (*n* = 281)	ND (*n* = 277)
QLQ-C30 scores (scale 0–100, high values are good)		
Global health status	71.3 (20.5)	71.0 (21.4)
Physical functioning	91.2 (15.6)	90.6 (16.7)
Role functioning	84.3 (26.3)	84.0 (25.7)
Emotional functioning	74.5 (22.3)	74.8 (22.6)
Cognitive functioning	85.1 (20.7)	84.8 (20.9)
Social functioning	82.5 (25.1)	80.3 (24.4)
QLQ-C30 symptoms and side effects (scale 0–100, low values are good)		
Fatigue	77.0 (24.1)	77.4 (23.6)
Nausea and vomiting	94.7 (13.4)	95.8 (11.6)
Pain	78.4 (26.2)	77.8 (26.1)
Dyspnoea	90.6 (18.6)	89.5 (22.3)
Insomnia	72.2 (30.7)	69.1 (31.3)
Appetite loss	81.5 (28.3)	83.5 (25.6)
Constipation	87.0 (23.4)	86.9 (23.2)
Diarrhoea	95.3 (13.7)	94.6 (14.9)
Financial difficulties	76.8 (33.9)	79.7 (30.7)

TABLE 42 Baseline quality-of-life scores (*continued*)

Quality-of-life scale	Trial arm, mean (SD)	
	Surveillance ($n = 281$)	ND ($n = 277$)
H&N35 scores (H&N quality of life) (scale 0–100, high values are good)		
Pain	75.7 (23.7)	73.5 (25.2)
Swallowing	87.0 (19.9)	85.1 (22.5)
Senses problems	91.6 (17.4)	90.6 (17.8)
Speech problems	87.8 (18.5)	87.5 (18.2)
Trouble with social eating	86.7 (22.2)	87.5 (21.8)
Trouble with social contact	93.0 (15.7)	93.9 (14.0)
Less sexuality	75.9 (32.3)	78.0 (32.9)
Teeth	87.1 (26.6)	85.5 (26.6)
Opening mouth	87.5 (25.8)	85.3 (28.0)
Dry mouth	77.4 (28.6)	81.7 (24.4)
Sticky saliva	84.0 (25.4)	84.4 (27.2)
Coughing	74.2 (25.4)	74.8 (26.9)
Felt ill	84.2 (25.5)	86.5 (21.0)
Used painkillers	36.4 (48.2)	33.8 (47.4)
Taken nutritional supplements	84.3 (36.5)	84.5 (36.3)
Used a feeding tube	97.3 (16.2)	94.6 (22.7)
Weight loss	72.7 (44.7)	71.4 (45.3)
Weight gain	85.0 (35.8)	85.5 (35.3)
MDADI (dysphagia scales) (scale 0–100, high values are good)		
Global	76.5 (29.0)	75.3 (29.7)
Emotional	77.8 (18.9)	76.8 (18.0)
Functional	81.2 (19.7)	80.8 (19.1)
Physical	76.2 (24.5)	75.4 (23.3)
Total	78.1 (20.3)	77.2 (19.4)

SD, standard deviation.

The European Organisation for Research and Treatment of Cancer's Quality of Life Questionnaire for Cancer (with 30 questions) quality-of-life outcomes

In *Tables 43–47*, 'mean treatment difference' is the mean change from baseline in the surveillance arm minus the mean change from baseline in the ND arm. For these preplanned analyses of quality-of-life scores, no *p*-value was specified for determining a significant difference. In *Table 43*, global and functioning scales are low = bad, high = good. So, a positive mean treatment difference is good for the PET–CT arm. There are no clear differences as, for example, in global health status depicted in *Figure 13*, but results are slightly more favourable for the surveillance arm. It suggests that the symptoms and quality-of-life effects of ND are less impactful that those of the CRT.

TABLE 43 The EORTC's QLQ-C30 quality-of-life treatment differences

Scales/measures	Trial arm, number of patients		Mean treatment difference	*p*-value (Wilcoxon)
	PET–CT	ND		
QLQ-C30 global health status				
Post CRT	196	186	0.12	0.97
6 months	192	169	4.94	0.03
12 months	182	171	3.03	0.09
24 months	180	166	–0.81	0.85
QLQ-C30 physical functioning				
Post CRT	198	190	–0.45	0.93
6 months	190	173	3.51	0.08
12 months	183	173	4.29	0.01
24 months	179	170	–0.9	0.89
QLQ-C30 role functioning				
Post CRT	196	187	0.81	0.93
6 months	188	172	3.71	0.14
12 months	182	169	12.09	0.001
24 months	177	169	–0.71	0.74
QLQ-C30 emotional functioning				
Post CRT	197	185	0.97	0.82
6 months	192	168	3.83	0.26
12 months	183	170	4.05	0.34
24 months	180	166	0.86	0.48
QLQ-C30 cognitive functioning				
Post CRT	194	181	–0.09	0.82
6 months	190	168	1.24	0.48
12 months	178	167	4.63	0.31
24 months	175	161	0.15	0.93
QLQ-C30 social functioning				
Post CRT	191	180	–2.69	0.64
6 months	184	166	2.51	0.41
12 months	181	168	5.37	0.07
24 months	176	163	–5.57	0.04

TABLE 44 The EORTC's QLQ-C30 symptom scale treatment differences

Scales/measures	Trial arm, number of patients		Mean treatment difference	*p*-value (Wilcoxon)
	PET–CT	ND		
QLQ-C30 fatigue				
Post CRT	199	190	–2.35	0.52
6 months	191	173	5.25	0.08
12 months	184	173	4.75	0.07
24 months	180	170	4.52	0.20
QLQ-C30 nausea and vomiting				
Post CRT	195	189	2.81	0.62
6 months	189	172	–1.48	0.29
12 months	184	171	3.09	0.06
24 months	180	167	–1.17	0.43
QLQ-C30 pain (i.e. general)				
Post CRT	191	182	0.46	0.92
6 months	185	165	7.65	0.01
12 months	179	169	8.49	0.01
24 months	176	165	3.98	0.11
QLQ-C30 dyspnoea				
Post CRT	197	188	0.91	0.69
6 months	188	173	0.3	0.60
12 months	183	171	3.79	0.08
24 months	179	170	1.34	0.69
QLQ-C30 insomnia				
Post CRT	197	190	–2.45	0.28
6 months	191	173	5.07	0.21
12 months	184	172	–1.1	0.93
24 months	179	170	–4.14	0.19
QLQ-C30 appetite loss				
Post CRT	195	189	6.42	0.10
6 months	189	172	7.22	0.11
12 months	184	172	7.49	0.09
24 months	179	167	4.35	0.35
QLQ-C30 constipation				
Post CRT	198	189	–4.48	0.28
6 months	191	172	–0.43	0.50
12 months	183	171	0.99	0.57
24 months	180	169	–0.33	0.70

continued

TABLE 44 The EORTC's QLQ-C30 symptom scale treatment differences (*continued*)

Scales/measures	Trial arm, number of patients: PET–CT	ND	Mean treatment difference	*p*-value (Wilcoxon)
QLQ-C30 diarrhoea				
Post CRT	198	186	–0.21	0.99
6 months	192	169	–3.27	0.13
12 months	182	170	–2.88	0.05
24 months	180	167	–0.17	0.94
QLQ-C30 financial difficulties				
Post CRT	195	183	–1.28	0.79
6 months	189	168	–0.15	0.87
12 months	181	170	4.24	0.28
24 months	178	165	–1.69	0.73

TABLE 45 The EORTC's H&N35 H&N quality-of-life score treatment differences

Scales/measures	Trial arm, number of patients: PET–CT	ND	Mean treatment difference	*p*-value (Wilcoxon)
H&N35 pain (mouth, jaw or throat)				
Post CRT	197	189	–1.48	0.77
6 months	190	172	–1.83	0.53
12 months	182	172	–1.43	0.86
24 months	180	168	–3.67	0.42
H&N35 swallowing				
Post CRT	188	185	–1.41	0.68
6 months	188	169	–0.68	0.69
12 months	182	173	–1.07	0.75
24 months	180	168	–3.08	0.43
H&N35 senses problems				
Post CRT	193	182	–1.17	0.89
6 months	186	168	–1.76	0.70
12 months	178	169	0.6	0.89
24 months	176	167	–4.08	0.06
H&N35 speech problems				
Post CRT	198	189	0.31	0.99
6 months	190	173	–0.33	0.84
12 months	182	172	1.23	0.75
24 months	179	167	–1.21	0.38

TABLE 45 The EORTC's H&N35 H&N quality-of-life score treatment differences (*continued*)

Scales/measures	Trial arm, number of patients		Mean treatment difference	*p*-value (Wilcoxon)
	PET–CT	ND		
H&N35 trouble with social eating				
Post CRT	173	178	–1.43	0.84
6 months	182	166	2.55	0.54
12 months	180	170	2.77	0.30
24 months	176	166	0.23	0.99
H&N35 trouble with social contact				
Post CRT	198	189	1.96	0.55
6 months	190	173	1.47	0.38
12 months	182	172	6.96	0.01
24 months	179	167	0.12	0.87
H&N35 less sexuality				
Post CRT	163	161	2.79	0.62
6 months	163	156	2.03	0.54
12 months	161	156	6.42	0.21
24 months	154	148	1.08	0.77
H&N35 problems with teeth				
Post CRT	191	178	0.09	0.86
6 months	186	168	3.12	0.21
12 months	179	169	6.92	0.10
24 months	176	164	–2.9	0.44
H&N35 problems opening mouth wide				
Post CRT	198	185	–6.13	0.23
6 months	191	172	0.92	0.59
12 months	181	172	2.93	0.29
24 months	179	167	5.75	0.08
H&N35 dry mouth				
Post CRT	198	188	–0.12	0.95
6 months	191	172	–0.73	0.66
12 months	181	171	7.38	0.08
24 months	179	168	–2.39	0.44
H&N35 sticky saliva				
Post CRT	196	189	0.49	0.95
6 months	188	172	–4.7	0.14
12 months	181	171	–4.69	0.23
24 months	177	168	–3.47	0.42

continued

TABLE 45 The EORTC's H&N35 H&N quality-of-life score treatment differences (*continued*)

Scales/measures	Trial arm, number of patients		Mean treatment difference	*p*-value (Wilcoxon)
	PET–CT	ND		
H&N35 coughing				
Post CRT	197	189	–4.35	0.20
6 months	192	171	3.08	0.73
12 months	182	172	2.77	0.65
24 months	180	167	–0.05	0.81
H&N35 felt ill				
Post CRT	197	189	4.04	0.26
6 months	190	172	5.21	0.07
12 months	182	173	4.91	0.19
24 months	180	167	1.63	0.55
H&N35 used painkillers (yes/no)				
Post CRT	198	188	–1.32	0.79
6 months	189	172	3.87	0.53
12 months	181	171	16.45	0.01
24 months	179	166	4.75	0.61
H&N35 taken nutritional supplements (yes/no)				
Post CRT	195	186	2	0.53
6 months	189	171	6.18	0.26
12 months	182	167	0.97	0.89
24 months	177	165	1.79	0.74
H&N35 used a feeding tube (yes/no)				
Post CRT	197	187	–0.04	0.93
6 months	189	171	1.34	0.82
12 months	180	169	–2.16	0.62
24 months	178	167	–2.17	0.40
H&N35 weight loss (yes/no)				
Post CRT	186	183	–3.5	0.61
6 months	185	163	–5.93	0.38
12 months	175	169	1.53	0.83
24 months	175	163	1.49	0.78
H&N35 weight gain (yes/no)				
Post CRT	180	173	0.56	0.91
6 months	181	156	7.27	0.21
12 months	167	166	1.32	0.88
24 months	172	159	10.48	0.10

TABLE 46 The MDADI dysphagia score treatment differences

Scales/measures	Trial arm, number of patients		Mean treatment difference	*p*-value (Wilcoxon)
	PET–CT	ND		
MDADI dysphagia global				
Post CRT	187	182	0.95	0.65
6 months	180	167	8.46	0.05
12 months	173	161	4.31	0.17
24 months	168	161	–0.27	0.92
MDADI dysphagia emotional				
Post CRT	185	179	–1.68	0.36
6 months	178	163	2.11	0.43
12 months	173	164	2.83	0.44
24 months	168	160	–1.05	0.91
MDADI dysphagia functional				
Post CRT	189	181	–1.39	0.54
6 months	186	166	1.46	0.48
12 months	175	166	1.22	0.76
24 months	172	162	–0.87	0.77
MDADI dysphagia physical				
Post CRT	185	177	0.53	0.87
6 months	178	161	0.27	0.97
12 months	173	164	1.38	0.70
24 months	168	159	–0.49	0.99
MDADI dysphagia total				
Post CRT	189	183	–0.98	0.56
6 months	183	166	1.62	0.46
12 months	175	166	1.75	0.58
24 months	171	162	–0.64	0.99

TABLE 47 The EQ-5D health status treatment differences

Scales/measures	Trial arm, number of patients		Mean treatment difference	*p*-value (Wilcoxon)
	PET–CT	ND		
EQ-5D overall health status				
Post CRT	187	174	0.04	0.20
6 months	181	166	0.04	0.09
12 months	173	161	0.07	0.007
24 months	175	156	0.02	0.21

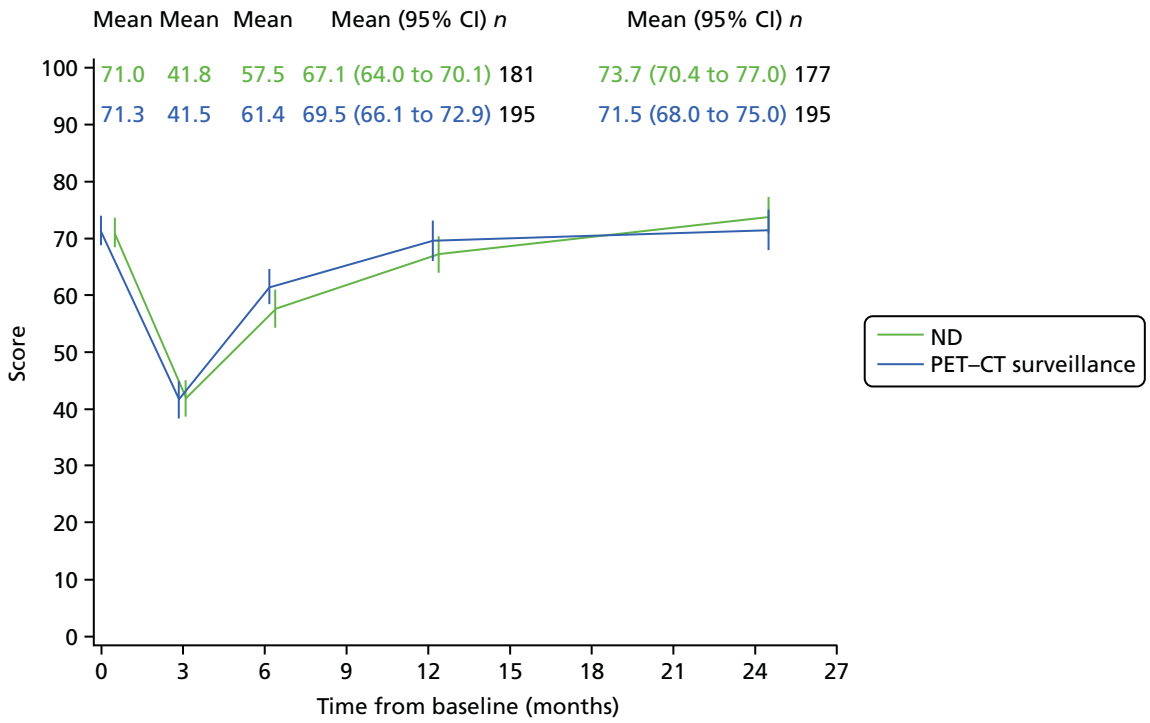


FIGURE 13 The EORTC's QLQ-C30 global health status. Reproduced from the New England Journal of Medicine, Mehanna H, Wong W-L, McConkey CC, Rahman JK, Robinson M, Hartley AGJ, PET-CT Surveillance versus Neck Dissection in Advanced Head and Neck Cancer, 374, 1444–54. Copyright © (2016) Massachusetts Medical Society. Reprinted with permission.[62]

The European Organisation for Research and Treatment of Cancer's Quality of Life Questionnaire for Cancer (with 30 questions) symptom scales

For simplicity of interpretation the symptom scales in *Table 44* have been transformed so that, here also, a positive mean treatment difference is good for the PET–CT arm. There are no clear differences, but results are slightly more favourable for the surveillance arm.

The European Organisation for Research and Treatment of Cancer's Quality of Life Questionnaire for Cancer head and neck module with 35 questions quality-of-life outcomes

All scales in *Table 45* have been set to low as bad and high as good, so a positive difference in mean treatment difference is good for the PET–CT arm. There is little to suggest differences between trial treatments. No difference is seen in pain in H&N cancer and in problems with swallowing (*Figures 14* and *15*). The difference in problems with teeth is small (*Figure 16*). Reported use of painkillers at 12 months was higher in the ND arm (*Figure 17*).

The MD Anderson Dysphagia Inventory dysphagia scales

All scales in *Table 46* have been set to low as bad and high as good, so a positive difference in mean treatment difference is good for the PET–CT arm. There is little to suggest differences between trial treatments (see, e.g., *Figure 18*).

European Quality of Life-5 Dimensions

The EQ-5D is used for the health resource study. In *Table 47*, positive values for mean treatment difference indicate a better response in the PET–CT arm.

Differences of at least 10% in quality-of-life scales

Another way of presenting these data is to display differences from baseline of at least 10%. *Figures 19–25* compare the randomised groups on that basis. *Tables 48–54* give similar information, comparing patients who underwent ND before CRT with those who did not. The swallowing measures in *Figures 21*, *24* and *25* show no difference between trial treatments, with both having a low proportion of improvers.

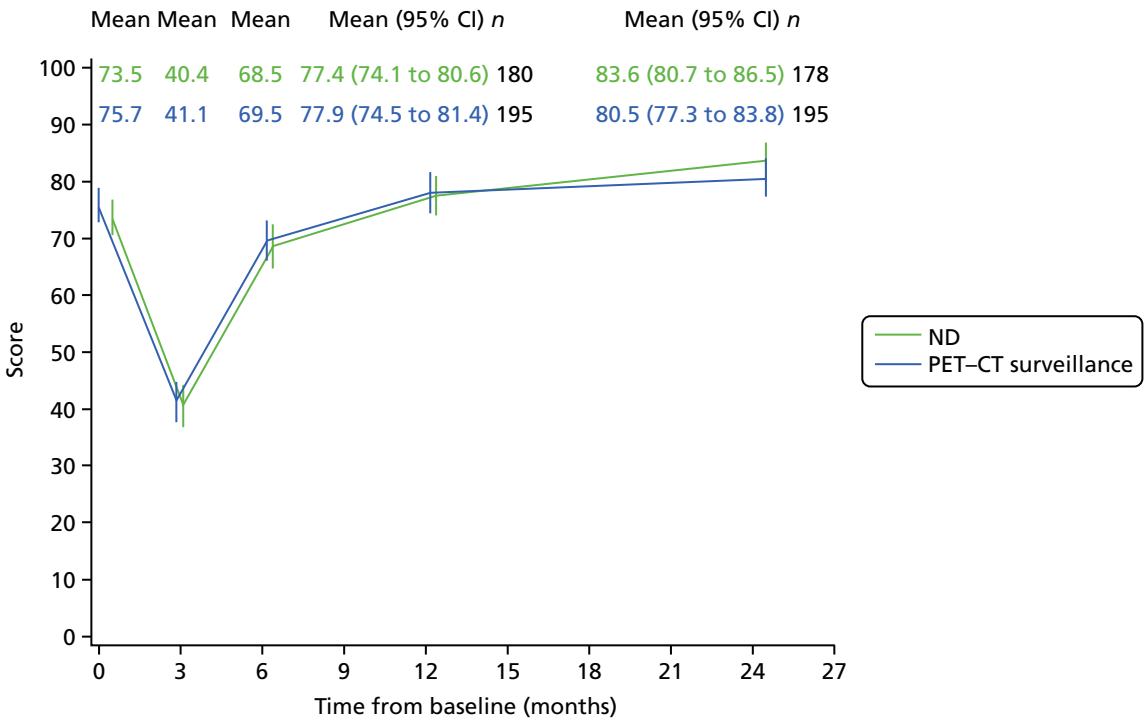


FIGURE 14 The EORTC's H&N35 pain in H&N cancer: mean scores unadjusted.

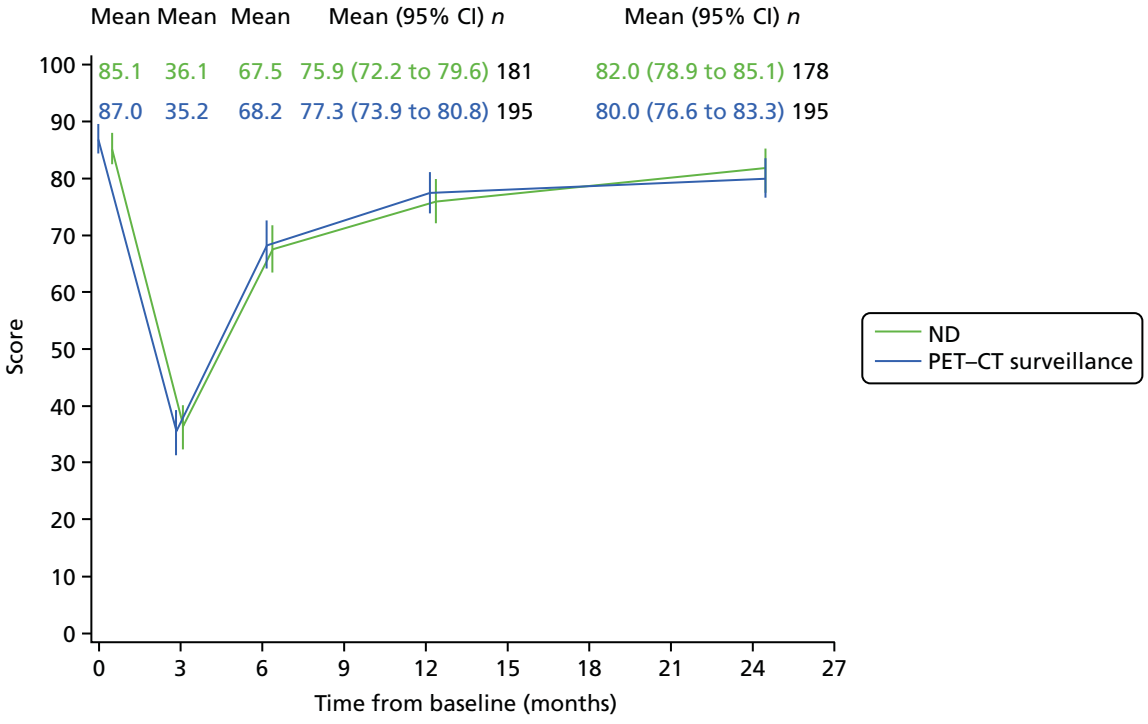


FIGURE 15 The EORTC's H&N35 problems with swallowing: mean scores unadjusted.

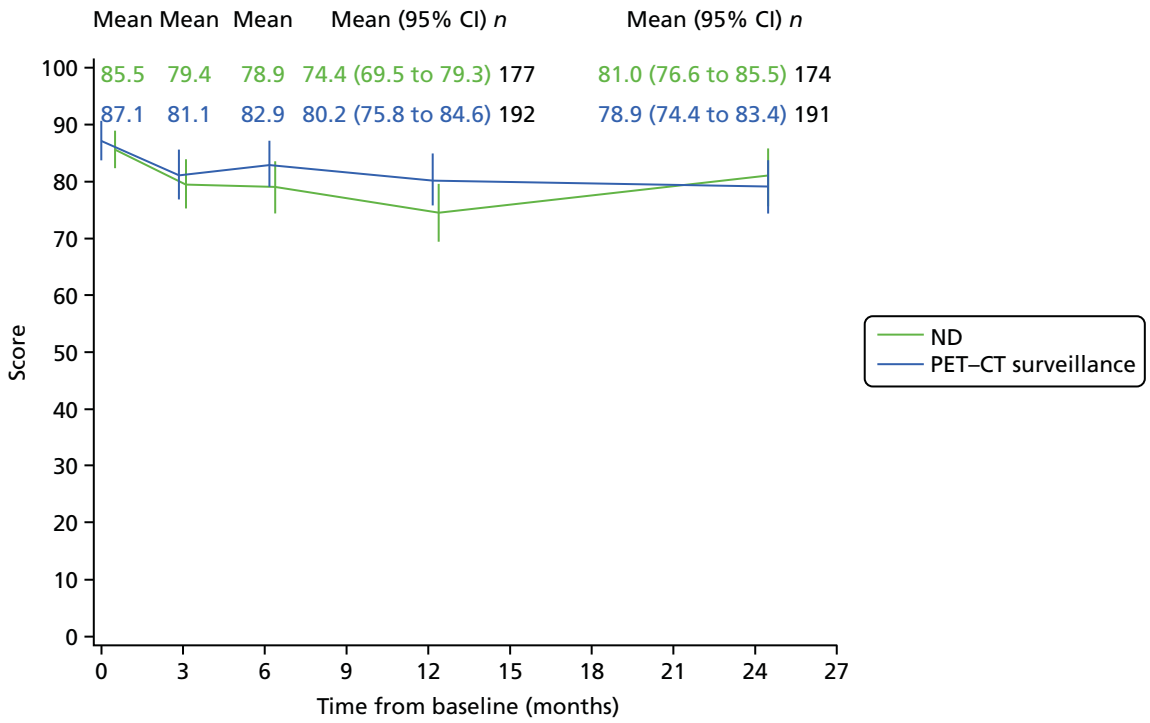


FIGURE 16 The EORTC's H&N35 problems with teeth: mean scores unadjusted.

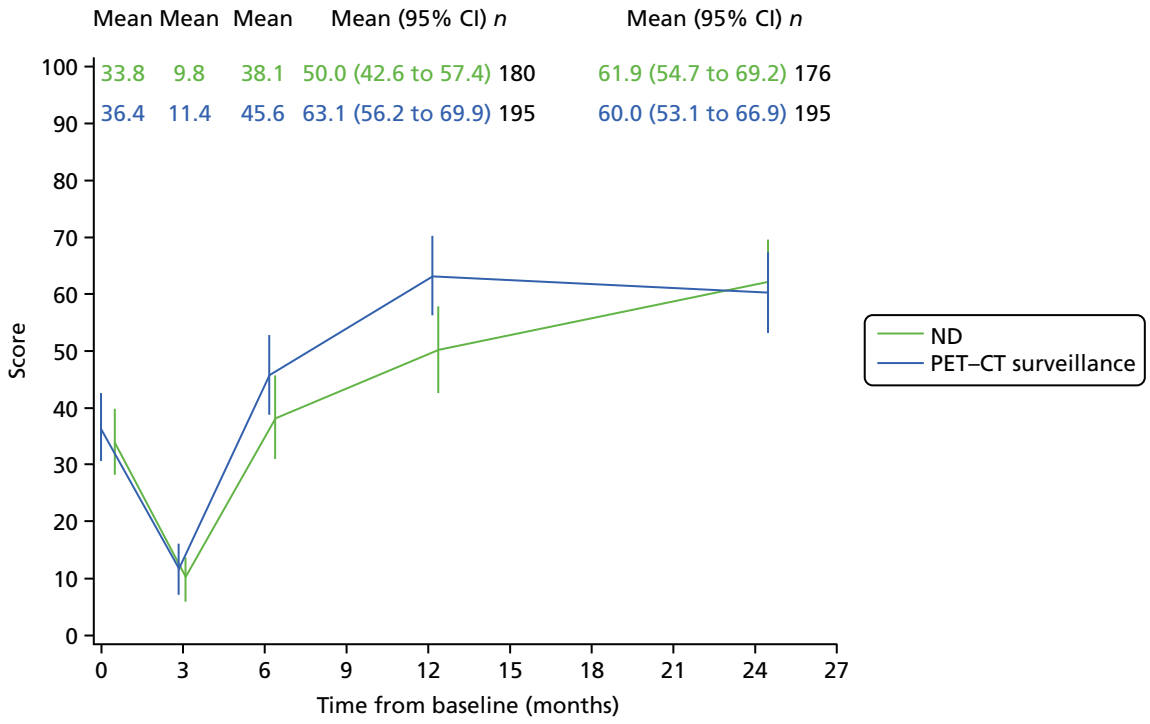


FIGURE 17 The EORTC's H&N35 used painkillers: mean scores unadjusted (higher score = less pain).

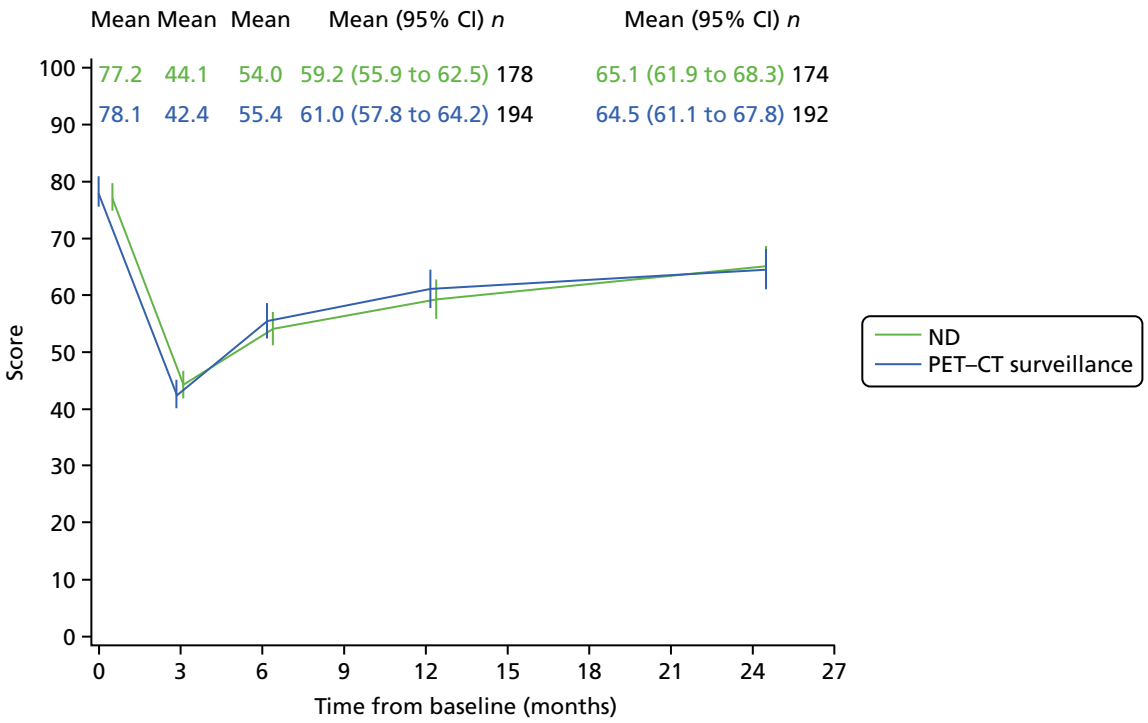


FIGURE 18 The MDADI dysphagia total: mean scores unadjusted.

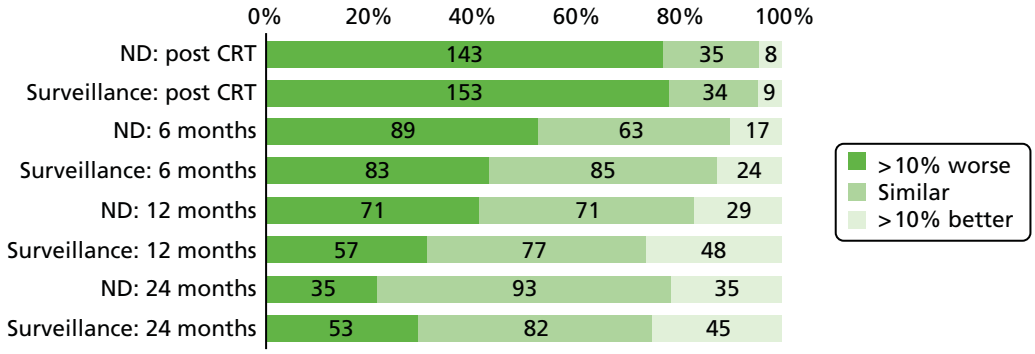


FIGURE 19 The EORTC's QLQ-C30 global health status, percentages changed by ≥ 10%.

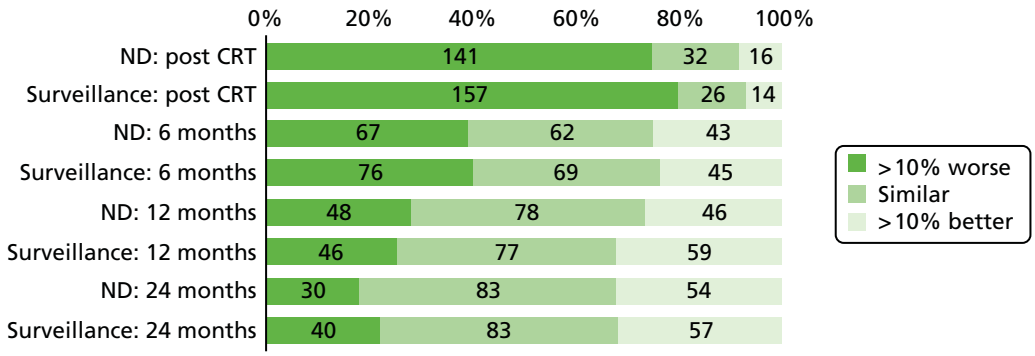


FIGURE 20 The EORTC's QLQ H&N35 pain scale, percentages changed by ≥ 10%.

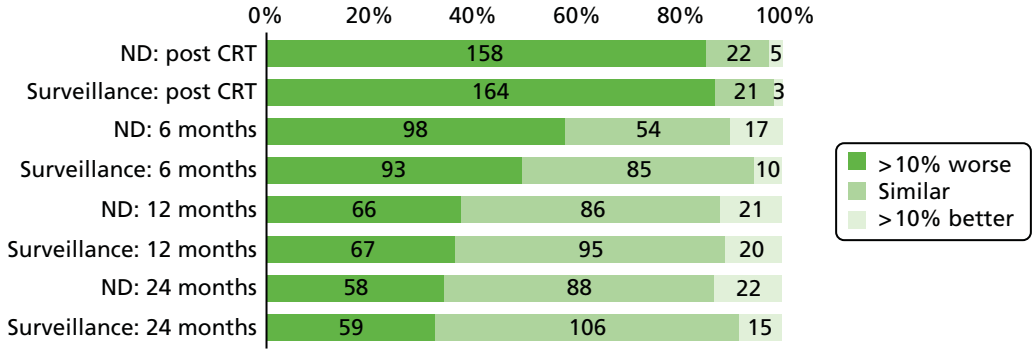


FIGURE 21 The EORTC's QLQ H&N35 swallowing scale, percentages changed by ≥ 10%.

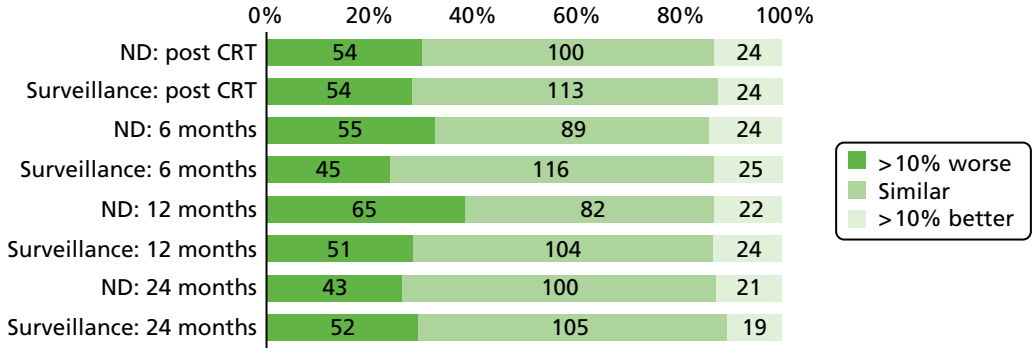


FIGURE 22 The EORTC's QLQ H&N35 problems with teeth, percentages changed by ≥ 10%.

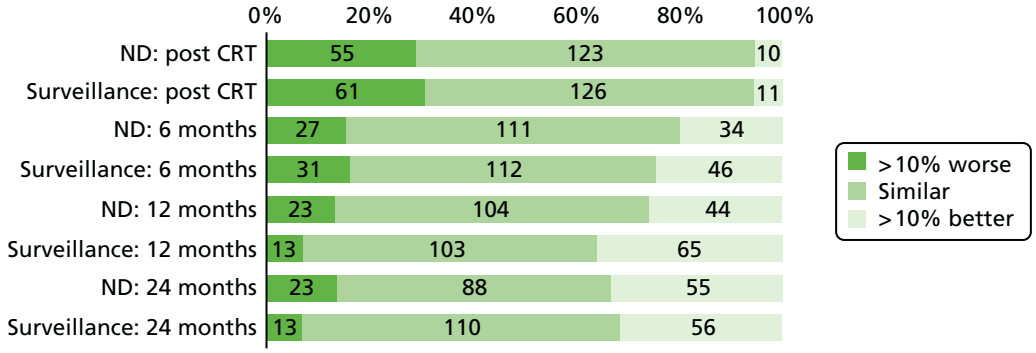


FIGURE 23 The EORTC's QLQ H&N35 use of painkillers, percentages changed by ≥ 10%.

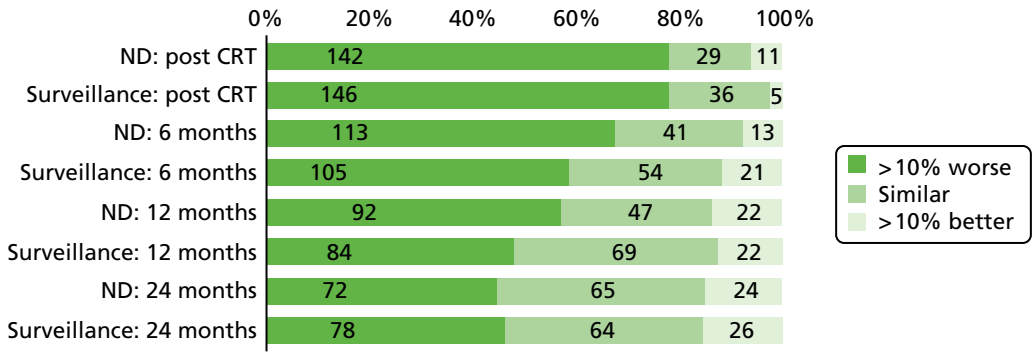


FIGURE 24 The MDADI dysphagia global scale, percentages changed by ≥ 10%.

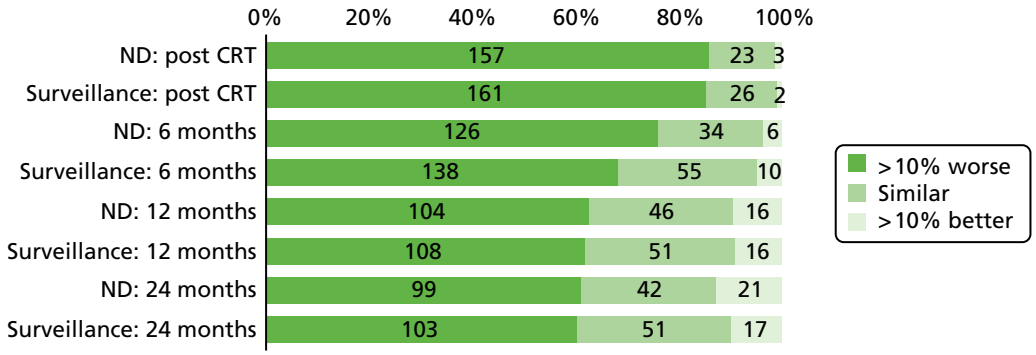


FIGURE 25 The MDADI dysphagia total scale, percentages changed by ≥ 10%.

TABLE 48 The EORTC's QLQ-C30 global health status: percentages changed by ≥ 10%, by actual timing of ND before or after CRT

Time point and group	Decrease by ≥ 10%	Similar	Increase by ≥ 10% (improvement)	Total
Post CRT*, n *(%)				
ND planned before CRT	42 (84.0)	6 (12.0)	2 (4.0)	50
All others	254 (76.5)	63 (19.0)	15 (4.5)	332
6 months*, n *(%)				
ND before CRT	28 (54.9)	17 (33.3)	6 (11.8)	51
All others	144 (46.5)	131 (42.3)	35 (11.3)	310
12 months*, n *(%)				
ND before CRT	21 (43.8)	17 (35.4)	10 (20.8)	48
All others	107 (35.1)	131 (43.0)	67 (22.0)	305
24 months*, n *(%)				
ND before CRT	14 (29.2)	25 (52.1)	9 (18.8)	48
All others	75 (25.2)	152 (51.0)	71 (23.8)	298

TABLE 49 The EORTC's H&N35 pain scale: percentages changed by ≥ 10%, by actual timing of ND before or after CRT

Time point and group	Decrease by ≥ 10%	Similar	Increase by ≥ 10% (improvement)	Total
Post CRT*, n *(%)				
ND planned before CRT	36 (72.0)	6 (12.0)	8 (16.0)	50
All others	262 (78.0)	52 (15.5)	22 (6.6)	336
6 months*, n *(%)				
ND planned before CRT	23 (44.2)	17 (32.7)	12 (23.1)	52
All others	120 (38.7)	114 (36.8)	76 (24.5)	310
12 months*, n *(%)				
ND planned before CRT	16 (33.3)	18 (37.5)	14 (29.2)	48
All others	782 (25.5)	137 (44.8)	91 (29.7)	306
24 months*, n *(%)				
ND planned before CRT	9 (18.8)	23 (47.9)	16 (33.3)	48
All others	62 (20.7)	143 (47.7)	95 (31.7)	300

TABLE 50 The EORTC's H&N35 swallowing symptoms scale: comparison by actual timing of ND for those who had a ND

Time point and group	Decrease by ≥ 10%	Similar	Increase by ≥ 10% (improvement)	Total
Post CRT*, n *(%)				
ND planned before CRT	39 (81.3)	8 (16.7)	1 (2.1)	48
All others	283 (87.1)	35 (10.8)	7 (2.2)	325
6 months*, n *(%)				
ND planned before CRT	27 (51.9)	18 (34.6)	7 (13.5)	52
All others	164 (53.8)	121 (39.7)	20 (6.6)	305
12 months*, n *(%)				
ND planned before CRT	20 (41.7)	21 (43.8)	7 (14.6)	48
All others	113 (36.8)	160 (52.1)	34 (11.1)	307
24 months*, n *(%)				
ND planned before CRT	17 (35.4)	25 (52.1)	6 (12.5)	48
All others	100 (33.3)	169 (56.3)	31 (10.3)	300

TABLE 51 The EORTC's QLQ-C30 problems with teeth scale: percentages changed by ≥ 10%, by actual timing of ND before or after CRT

Time point and group	Decrease by ≥ 10%	Similar	Increase by ≥ 10% (improvement)	Total
Post CRT*, n *(%)				
ND planned before CRT	17 (37.8)	26 (57.8)	2 (4.4)	45
All others	91 (28.1)	187 (57.7)	46 (14.2)	324
6 months*, n *(%)				
ND planned before CRT	16 (32.0)	29 (58.0)	5 (10.0)	50
All others	84 (27.6)	176 (57.9)	44 (14.5)	304
12 months*, n *(%)				
ND planned before CRT	15 (33.3)	24 (53.3)	6 (13.3)	45
All others	101 (33.3)	162 (53.5)	40 (13.2)	303
24 months*, n *(%)				
ND planned before CRT	14 (30.4)	29 (63.0)	3 (6.5)	46
All others	81 (27.6)	176 (59.9)	37 (12.6)	294

TABLE 52 The EORTC's QLQ-C30 used painkillers scale: percentages changed by ≥ 10%, by actual timing of ND before or after CRT

Time point and group	Decrease by ≥ 10%	Similar	Increase by ≥ 10% (improvement)	Total
Post CRT*, n *(%)				
ND planned before CRT	24 (48.0)	25 (50.0)	1 (2.0)	50
All others	92 (27.4)	224 (66.7)	20 (5.6)	336
6 months*, n *(%)				
ND planned before CRT	13 (24.5)	34 (64.2)	6 (11.3)	53
All others	45 (14.6)	189 (61.4)	74 (24.0)	308
12 months*, n *(%)				
ND planned before CRT	12 (25.0)	33 (68.8)	3 (6.3)	48
All others	24 (7.9)	174 (57.2)	106 (34.9)	304
24 months*, n *(%)				
ND planned before CRT	15 (31.3)	25 (52.1)	8 (16.7)	48
All others	21 (7.1)	173 (58.3)	103 (34.7)	297

TABLE 53 The MDADI dysphagia global scale: percentages changed by ≥ 10%, by actual timing of ND before or after CRT

Time point and group	Decrease by ≥ 10%	Similar	Increase by ≥ 10% (improvement)	Total
Post CRT*, n *(%)				
ND planned before CRT	35 (74.5)	9 (19.2)	3 (6.4)	47
All others	253 (78.6)	56 (17.4)	13 (4.0)	322
6 months*, n *(%)				
ND planned before CRT	32 (65.3)	15 (30.6)	2 (40.1)	49
All others	186 (62.4)	80 (26.9)	32 (10.7)	298
12 months*, n *(%)				
ND planned before CRT	21 (47.7)	17 (38.6)	6 (13.6)	44
All others	155 (53.5)	99 (34.1)	36 (12.4)	290
24 months*, n *(%)				
ND planned before CRT	20 (45.5)	17 (38.6)	7 (15.9)	44
All others	130 (45.6)	112 (39.3)	43 (15.1)	285

TABLE 54 The MDADI dysphagia total scale: percentages changed by ≥ 10%, by actual timing of ND before or after CRT

Time point and group	Decrease by ≥ 10%	Similar	Increase by ≥ 10% (improvement)	Total
Post CRT*, n *(%)				
ND planned before CRT	43 (89.6)	5 (10.4)	0 (0.0)	48
All others	275 (84.9)	44 (13.6)	5 (1.5)	324
6 months*, n *(%)				
ND planned before CRT	35 (70.0)	14 (28.0)	1 (2.0)	50
All others	229 (76.6)	55 (18.4)	15 (5.0)	299
12 months*, n *(%)				
ND planned before CRT	31 (67.4)	12 (26.1)	3 (6.5)	46
All others	181 (61.4)	85 (28.8)	29 (9.8)	295
24 months*, n *(%)				
ND planned before CRT	31 (68.9)	10 (22.2)	4 (8.9)	45
All others	171 (59.4)	83 (28.8)	34 (11.8)	288

A slightly greater percentage of patients who had an early ND had a decreased global health status. This is not generally the case for the more specific symptom scores in *Figures 20–25*.

p16 subgroups

The baseline scores are better in the p16-positive group than in the negative group. No difference is seen between treatment groups, but the plots in *Figures 26* and *27* suggest that the p16-positive group is affected more by CRT than the p16-negative group.

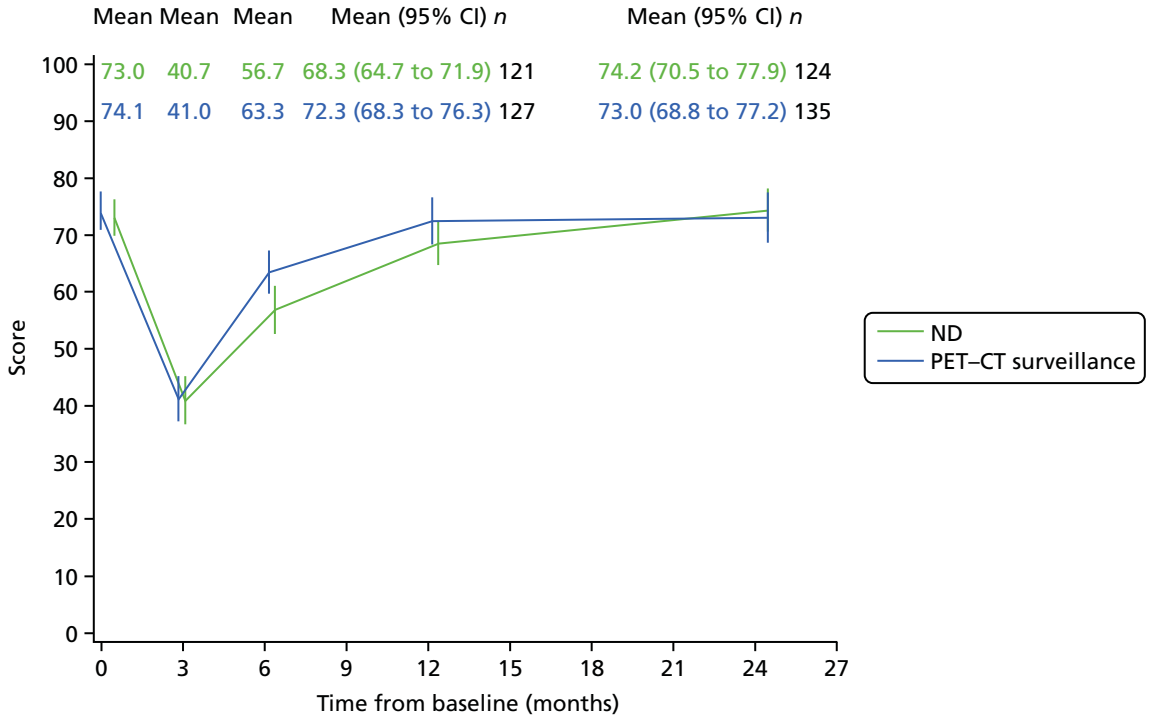


FIGURE 26 The EORTC's QLQ-C30 global health status: p16-positive group.

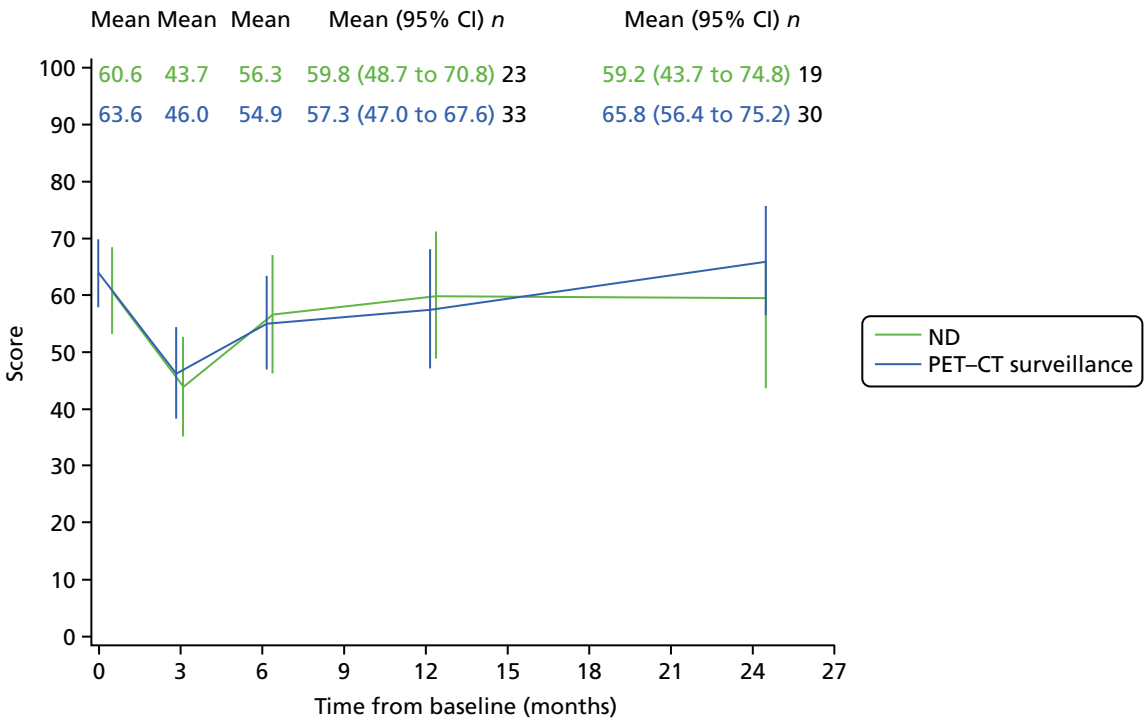


FIGURE 27 The EORTC's QLQ-C30 global health status: p16-negative group.

PET–CT scans and concordance

The PET–CT surveillance arm of the trial contained 282 patients, of whom 276 received CRT. Six patients did not receive PET–CT imaging: two died early, one progressed, two proceeded directly to ND and one missed an appointment and died before the rearranged date. PET–CT was performed in 270 patients, and the resulting assessments are summarised in *Table 55*.

Concordance rate for nodal disease

The concordance rate, on the basis of the readings reported by the local sites, is 247 out of 270, or 91.5%, but on review and a correct reading of the local report, the last group in the table (3 + 4 + 6 + 2 + 1) should be interpreted as concordant, giving a rate of 263 out of 270, or 97.4%. In total, 243 patients are concordant in primary tumour and neck nodes, 20 are concordant in neck nodes only and seven are discordant in neck nodes. Therefore, the concordance rate for the sites and the central laboratory is 97.4%, with a kappa value of 0.78 (95% CI 0.69 to 0.87). The summary assessments are related to NDs performed in *Table 56*.

Concordance in primary tumour

Similarly, for primary tumour the concordance rate for the sites and the central laboratory is 243 out of 266 or 91.4%, with a kappa value of 0.63 (95% CI 0.49 to 0.77).

True-negative rate

The number of patients in whom no neck cancer was seen on PET–CT on review and for whom no ND was carried out was 191. Among this group, 26 experienced recurrence, giving a true-negative rate of 86.4%. Only one recurrence was in the neck alone. Similarly, the true-negative rate for primary tumour was 29 out of 192 (84.9%).

TABLE 55 Positron emission tomography–computerised tomography concordance of local and central review assessments

Assessment					
Local		Central review			
Response in primary tumour	Response in neck nodes	Response in primary tumour	Response in neck nodes	Concordance	*n*
Concordant in primary tumour and nodes					
CR	CR	CR	CR	Yes	169
CR	Not CR	CR	Not CR	Yes	36
Not CR	CR	Not CR	CR	Yes	8
Not CR	Not CR	Not CR	Not CR	Yes	13
N/A	CR	N/A	CR	Yes	2
N/A	CR	CR	CR	Yes	2
Concordant in nodes only					
CR	Not CR	Not CR	Not CR	Yes	4
Not CR	Not CR	CR	Not CR	Yes	5
CR	CR	Not CR	CR	Yes	3
Not CR	CR	CR	CR	Yes	5
Discordant					
CR	CR	Not CR	Not CR	No	3
CR	CR	CR	Not CR	No	3
CR	Not CR	CR	CR	No	1
Discordant summary (local radiologist), concordant summary (central reviewer)					
CR	CR	Not CR	Not CR	No	3[a]
CR	CR	CR	Not CR	No	4[b]
CR	Not CR	CR	CR	No	6[c]
Not CR	CR	Not CR	Not CR	No	2[d]
Not CR	Not CR	Not CR	CR	No	1[c]

N/A, not applicable.
a Reviewed: text of local report expresses partial in nodes, making these concordant for nodes.
b Reviewed: text of local report expresses partial in nodes, making these concordant for primary tumour and nodes.
c Reviewed: text of local report expresses CR in nodes, making this concordant in primary tumour and nodes.
d Reviewed: text of local report expresses partial in nodes, making these concordant for primary tumour and nodes.

TABLE 56 Summary of PET–CT scan results and NDs performed

CR in			
Primary tumour	Neck nodes	ND	Number
Yes	Yes	No	181
Yes	Yes	Yes	4
Yes	No	Yes	36
Yes	No	No	11[a]
No	Yes	Yes	–
No	Yes	No	15
No	No	Yes	12[b]
No	No	No	7
Missing	Yes	No	4

a Reasons for the 11 having no ND when there was a CR in primary tumour, not in neck nodes: new lung primary tumour; disease progression or new primary lesion; liver metastases found at assessment; retroperitoneal nodes; recurrence confirmed 5 weeks after post-CRT assessment; biopsy confirmed no malignancy; biopsy – no malignancy remains, only debris; neck node judged by MDT to be reactive; faint activity in lymph node on PET–CT scan, but deemed to be a CR; PET scan reviewed at MDT and considered negative; and judged to have complete metabolic response.

b Reasons for ND when there was not a CR in the primary tumour and not in neck nodes: three were CRs in the primary tumour on review. There is no further information on the rest.

Chapter 4 Economic evaluation

Introduction

An economic evaluation was conducted to assess the cost-effectiveness of the PET–CT-guided watch-and-wait policy compared with planned ND. The evaluation consisted of two components: (1) a WT analysis, in which cost-effectiveness was assessed over the 24-month trial period using individual patient data collected in the trial; and (2) a decision-analytic model analysis, in which cost-effectiveness was assessed over a lifetime horizon, using standard modelling techniques applied to the trial data in order to extrapolate the trial results. The primary analysis was conducted from a NHS secondary care perspective (i.e. including NHS hospital costs). In addition, sensitivity analyses were conducted using (1) a NHS and Personal Social Services (PSS) perspective and (2) a societal perspective.

Methods

Within-trial analysis

Individual patient data collected in the trial were used to determine the costs and quality-adjusted life-years (QALYs) associated with each treatment arm. QALYs were derived using patient EQ-5D questionnaire responses, and costs were calculated using information collected in case report forms and patient- and carer-reported questionnaires. Cost-effectiveness was assessed as the incremental cost-effectiveness ratio (ICER) and incremental net health benefit (INB). Future costs and health outcomes (beyond 1 year) were discounted at an annual rate of 3.5%, as per the National Institute for Health and Care Excellence (NICE)'s *Guide to the Methods of Technology Appraisal 2013*.[63]

Quality-adjusted life-years

In the economic evaluation [both WT and lifetime decision model (DM)] patient health benefit was measured in terms of QALYs. QALYs provide a generic measure of overall patient health and are a composite measure of patient survival weighted by quality of life (utility) over time; for example, 1 year in full health is equivalent to 1 QALY, whereas 1 year at half of full health is equivalent to 0.5 QALYs. Expression of health benefit in terms of QALYs allows decision-makers to make a direct comparison of the cost-effectiveness of interventions across different disease areas and indications, and NICE currently recommends the use of QALYs in cost-effectiveness analyses in its reference case.[63]

Information on patient health-related quality of life during the trial was collected using patient responses to the EQ-5D-3L (three levels) questionnaire.[64] The EQ-5D questionnaire asks patients to identify their current health status across five health domains: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Patients can respond that they currently have no problems, some problems or major problems within each of the given domains; these responses are coded as 1, 2 or 3, respectively. According to their responses to the EQ-5D questionnaire, patients may therefore be classified in one of 243 (3^5) health states.

Patients in the PET-NECK trial were asked to complete the EQ-5D questionnaire at baseline, during treatment (2 weeks post CRT) and at 3, 6, 12 and 24 months post randomisation. For each time point patients' identified health state classifications were converted into quality-of-life values by applying the standard European Quality of Life UK tariff values;[65] these values represent the UK general public's preference for each of the possible EQ-5D-defined health states and provide a utility value for each of the health states. For each time point, patient utilities across the two arms were compared using the Wilcoxon signed-rank test.[66]

Quality-adjusted life-years were calculated by combining patient utility values with OS data. The Kaplan–Meier method was used to account for loss to follow-up and life-years were calculated by taking the area under the survival curve. The area under the survival curve over each time period was then weighted by the corresponding utility to calculate patient QALYs.[67]

Resource use and costs

Data on patient and carer consumption of NHS, PSS and personal resources during the trial were obtained from a combination of case report forms (completed by nurses/clinicians at the enrolled hospital) and patient- and carer-reported questionnaires. Case report forms were completed throughout the trial for all participants to collect information on resource use related to a range of secondary care activities, including node dissection/salvage surgery, radiotherapy, chemotherapy, SAEs, patient follow-up assessments and recurrence events. Patient report forms were used for a subset of patients ($n = 42$) to collect additional data on secondary care activity outside the patient's enrolled hospital oncology department, primary and community care activity (such as GP visits, nurse visits, counselling and therapy services, carers and social worker visits) and patient societal costs (i.e. patient travel expenses, equipment costs, one-off expenses and lost earnings resulting from illness). For this subset of patients, carer report forms were also used for those patients with a friend or family member acting as an informal carer, in order to identify carer societal costs (i.e. carer travel expenses, equipment costs, one-off expenses and lost earnings resulting from the associated patient's illness). In order to ensure quality of reported data and to minimise the burden of reporting on patients and carers, patient and carer report forms were collected for patients enrolled at the two highest-recruiting sites only [University Hospital Birmingham (UHB) and UHCW]. Individuals were asked to recall their use of services over the previous 3-month period (or since completion of the last form when appropriate), at the same time points as for the EQ-5D questionnaire (i.e. at baseline, treatment and 3, 6, 12 and 24 months post randomisation).

Owing to the small sample numbers of patients and carers providing self-reported cost data (42 patients and 35 carers) and to the significant uncertainty involved in imputing these data for the whole trial population, these data were excluded from the base-case analysis of costs. The base-case analysis includes data on secondary care resource use collected in the case report forms only and, therefore, adopts a NHS secondary care perspective. Sensitivity analyses were conducted in order to assess the impact of (1) including the additional patient-reported NHS and PSS data (i.e. using a NHS and PSS perspective) and (2) including all of the additional patient- and carer-reported data (i.e. using a societal perspective).

Costs were estimated by combining the resource use data with unit costs (*Table 57*) obtained from national sources including NHS national reference costs[69] and the Personal Social Services Research Unit (PSSRU)[70] costs for NHS and PSS service use costs, and the *British National Formulary*[72] and *Drugs and Pharmaceutical Electronic Market Information (eMit)*[71] for medication costs. The results of the economic evaluation are reported in pounds sterling (price year 2015). Unit costs reported in 2013/14 prices were inflated to 2015 prices in the analysis using the consumer price index inflation value reported by the Office for National Statistics (note that the prices reported in the original source were non-inflated).[76] Costs were compared between the two arms using the Wilcoxon signed-rank test.[66]

Missing data

For the calculation of QALYs, multiple imputation using chained equations was conducted using the 'mice' package in R (The R Foundation for Statistical Computing, Vienna, Austria). Data on patient baseline cancer stage, sex and age were used to impute missing EQ-5D data across the data collection time points. A sensitivity analysis was conducted based on a complete-case analysis in order to explore the impact of using imputation (see *Within-trial analysis*, *Uncertainty and sensitivity analyses*).

For costs, data on societal, primary and community care resource use and secondary care resource use outside the enrolled hospitals were collected for only a small subset of the overall trial population using patient- and carer-reported questionnaires. As the sample size of this subpopulation was small (patients, $n = 42$; carers, $n = 35$), the base-case analysis was conducted using costs derived from the case report forms (collected for the whole trial population and with complete data) only. Sensitivity analyses were

TABLE 57 Unit costs applied to resource use items in the WT analysis

Resource item	Unit cost	Source	Details/assumptions
Secondary care costs (relating to case report form items)			
Inpatient 'hotel cost' (per night cost)	£200	East and North Hertfordshire NHS Trust's *Performance Report Month 11*[68]	
Oncology assessment	£181	*Reference Costs 2013–2014*[69]	Medical oncology first face-to-face attendance. CC: WF01B. SC: 370
Cardiology assessment	£160	*Reference Costs 2013–2014*[69]	Cardiology first face-to-face attendance. CC: WF01B. SC: 320
Respiratory assessment	£186	*Reference Costs 2013–2014*[69]	Respiratory medicine first face-to-face attendance. CC: WF01B. SC: 340
Other assessment	£196	*Reference Costs 2013–2014*[69]	General medicine first face-to-face attendance. CC: WF01B. SC: 300
Dental assessment	£126	*Reference Costs 2013–2014*[69]	Dental medicine first face-to-face attendance. CC: WF01B. SC: 450
Nasopharyngoscopy	£114	*Reference Costs 2013–2014*[69]	Diagnostic nasopharyngoscopy, aged ≥ 19 years. CC: CA71A. SC: 370
Fine-needle aspiration	£164	*Reference Costs 2013–2014*[69]	Minor maxillofacial procedures. CC: CA95Z. SC: 144
Surgery assessment	£150	*Reference Costs 2013–2014*[69]	General surgery first face-to-face attendance. CC: WF01B. SC: 100
CT scan	£147	*Reference Costs 2013–2014*[69]	CT scan, more than three areas. CC: RA14Z. SC: DIAGIMOP
PET–CT scan	£649	*Reference Costs 2013–2014*[69]	Nuclear medicine, category 8 (PET–CT). CC: RA42Z
MRI scan	£145	*Reference Costs 2013–2014*[69]	MRI scan, one area, post contrast only, aged ≥ 19 years. CC: RA02A. SC: DIAGIMOP
Radiography	£40	Personal communication with LTHT (Dr Peter Hall, University of Leeds, 2015, personal communication)	
Ultrasound	£76	*Reference Costs 2013–2014*[69]	Ultrasound mobile scan or intraoperative procedures, < 20 minutes. CC: RA25Z. SC: DIAGIMOP
Other radiography assessment	£88	*Reference Costs 2013–2014*[69]	Clinical oncology (previously radiotherapy) first attendance. CC: WF01D. SC: 800
Nurse assessment	£100	*Unit Costs of Health and Social Care*[70]	Assume an assessment is equivalent to 1 hour of contact time
Palliative care assessment	£97	*Unit Costs of Health and Social Care*[70]	
Social other	£37	*Unit Costs of Health and Social Care*[70]	Assume equivalent to speech/diet assessment
Speech assessment	£37	*Unit Costs of Health and Social Care*[70]	Assume an assessment is equivalent to 1 hour of contact time
Dietitian assessment	£37	*Unit Costs of Health and Social Care*[70]	As above

continued

TABLE 57 Unit costs applied to resource use items in the WT analysis (*continued*)

Resource item	Unit cost	Source	Details/assumptions
Rehabilitation assessment	£36	*Unit Costs of Health and Social Care*[70]	Assume equivalent to 1 hour of occupational therapist contact time
Psychology assessment	£138	*Unit Costs of Health and Social Care*[70]	Assume hospital assessment cost equivalent to community visit cost
Counselling assessment	£50	*Unit Costs of Health and Social Care*[70]	As above
Peg swab	£7	*Reference Costs 2013–2014*[69]	Directly accessed – pathology services. Microbiology. CC: DAPS07
Bloods: haematology	£3	*Reference Costs 2013–2014*[69]	Directly accessed – pathology services. Haematology. CC: DAPS05
Bloods: biochemistry	£1	*Reference Costs 2013–2014*[69]	Directly accessed – pathology services. Clinical biochemistry. CC: DAPS04
Bloods: microbiology	£7	*Reference Costs 2013–2014*[69]	Directly accessed – pathology services. Microbiology. CC: DAPS07
Bloods: other	£8	*Reference Costs 2013–2014*[69]	Directly accessed – pathology services. Other. CC: DAPS09
Secondary care costs (relating to patient-reported items)			
Short-stay (≤ 2 days) inpatient cost	£611	*Unit Costs of Health and Social Care*[70]	
Long-stay (> 2 days) inpatient cost	£2716	*Unit Costs of Health and Social Care*[70]	
Hospital day centre	£119	*Reference Costs 2013–2014*[69]	Inpatient specialist palliative care, same day, aged ≥ 19 years and over. SC: DCRDN. CC: SD02A
Outpatient visit	£109	*Unit Costs of Health and Social Care*[70]	
Accident and emergency visit	£135	*Reference Costs 2013–2014*[69]	Total outpatient attendances. SC: 180
Nursing/convalescent home	£82	*Unit Costs of Health and Social Care*[70]	Assume cost for 1 day and night equals the reported private-sector nursing home cost per week/7
Primary and community care service costs (relating to patient-reported items)			
GP surgery visit (telephone call)	£46 (28)	*Unit Costs of Health and Social Care*[70]	
GP home visit	£67	*Unit Costs of Health and Social Care*[70]	Assume equal to reported cost for 17-minute surgery visit
District nurse home visit (telephone call)	£66 (£11)	*Unit Costs of Health and Social Care*[70]	Assume each visit equal to 1 hour of contact time and a call is equivalent to 10 minutes of contact time
Social worker visit (telephone call)	£79 (£13)	*Unit Costs of Health and Social Care*[70]	As above
Physiotherapist visit (telephone call)	£36 (£6)	*Unit Costs of Health and Social Care*[70]	As above
Occupational therapist visit (telephone call)	£36 (£6)	*Unit Costs of Health and Social Care*[70]	As above
Counsellor visit (telephone call)	£50 (£8)	*Unit Costs of Health and Social Care*[70]	As above

TABLE 57 Unit costs applied to resource use items in the WT analysis (*continued*)

Resource item	Unit cost	Source	Details/assumptions
Home help service	£24 (£4)	*Unit Costs of Health and Social Care*[70]	Assume a visit is equal to 1 hour of weekday contact and a call is equivalent to 10 minutes of this time
Psychiatrist	£138	*Unit Costs of Health and Social Care*[70]	Assume a visit is equal to 1 hour of contact time and a call is equivalent to 10 minutes of contact time
Day centre	£24	*Unit Costs of Health and Social Care*[70]	Assume equivalent to home help service visit
Chemotherapy drug costs			
5-fluorouracil	£3.47	*Drugs and pharmaceutical electronic market information (eMit)*, 2015[71]	5 g/100-ml vial, 5%, size 1
Cisplatin	£16.69	*Drugs and pharmaceutical electronic market information (eMit)*, 2015[71]	100 mg/100-ml vial
Carboplatin	£28.89	*Drugs and pharmaceutical electronic market information (eMit)*, 2015[71]	600 mg/60-ml vial
Cetuximab	£890.50	*British National Formulary*, 2015[72]	5 mg/ml, 100-ml vial
Docetaxel	£29.78	*Drugs and pharmaceutical electronic market information (eMit)*, 2015[71]	160 mg/8-ml vial
Delivery cost	£328	*Reference Costs 2013–2014*[69]	Deliver subsequent elements of a chemotherapy cycle. CC: SB15Z. SC: DCRDN
Radiotherapy costs			
Radiotherapy delivery	£149	*Reference Costs 2013–2014*[69]	Deliver a fraction of complex treatment on a megavoltage machine. CC: SC23Z. SC: DCRDN
Radiotherapy planning visit	£1587	*Reference Costs 2013–2014*[69]	Preparation for IMART, with technology support. CC: SC41Z. SC: DCRDN
Surgery costs			
Node dissection	£3548	*Reference Costs 2013–2014*[69]	Elective inpatient. Intermediate maxillofacial procedures. CC: CA94Z
Salvage surgery	£7722	*Reference Costs 2013–2014*[69]	Elective inpatient. Major maxillofacial procedures, aged ≥ 19 years, with a complexity and comorbidity score of ≥ 1. CC: CA93A
Follow-up visit assessment costs			
Anaesthetic examination	£85	*Reference Costs 2013–2014*[69]	Anaesthetics: diagnostic, laryngoscopy or pharyngoscopy, aged ≥ 19 years. CC: CA69A
Biopsy	£164	*Reference Costs 2013–2014*[69]	Maxillofacial surgery: minor maxillofacial procedures. CC: CA95Z
Clinical examination	£109	*Reference Costs 2013–2014*[69]	Maxillofacial surgery: Diagnostic, laryngoscopy or pharyngoscopy, aged ≥ 19 years. CC: CA69A

continued

TABLE 57 Unit costs applied to resource use items in the WT analysis (*continued*)

Resource item	Unit cost	Source	Details/assumptions
Recurrence treatment costs			
Brachytherapy	£2393	*Reference Costs 2013–2014*[69]	1 × preparation for interstitial brachytherapy (£1196). CC: DCRDN. SC: SC55Z 1 × deliver a fraction of intraluminal brachytherapy (£1197). CC: DCRDN. SC: SC30Z
Chemotherapy course	£4753	*Reference Costs 2013–2014*[69]	Six × procure chemotherapy drugs for regimens in band 2 (£323). CC: DCRDN. SC: SB02Z 1 × deliver more complex parenteral chemotherapy at first attendance (£317). CC: DCRDN. SC: SB13Z 5 × deliver subsequent elements of a chemotherapy cycle (£328). CC: DCRDN. SC: SB15Z 6 × medical oncology follow-up. CC: WF01A. SC: 370
Radiotherapy course	£1744		Assumes 1 radiotherapy planning visit and 1 delivery
Annual supportive care cost post distant recurrence	£1682	Hall *et al.*, 2014[73]	
Terminal month palliative care cost post distant recurrence	£1051	Hall *et al.*, 2014[73]	
Resection/free flap reconstruction	£7722	*Reference Costs 2013–2014*[69]	Elective inpatient. Major maxillofacial procedures, aged ≥ 19 years , with a complexity and comorbidity score of ≥ 1. CC: CA93A
Societal costs (relating to patient- and carer-reported forms)			
Average wage per hour	£15.11	Office for National Statistics' *Annual Survey of Hours and Earnings, 2014 Provisional Results*[74]	
Cost per mile of car travel	£0.67	NHS reimbursement data[75]	

CC, currency code; LTHT, Leeds Teaching Hospitals NHS Trust; SC, service code.

Note

The cost of MDT discussions was not included as this information was not collected on the trial case report forms and, therefore, there are no activity data on which to base this cost. Based on consultation with clinicians, it is unlikely that there would be a difference in the cost of MDT across the arms, as all patients would require a MDT discussion regardless of eventual treatment.

conducted, which included the additional resource use items reported in the patient- and carer-reported questionnaires. These forms were relatively complete, with most patients returning their forms at each of the follow-up time points (*Table 58*). In the case of missing data within each of the forms, values were based on imputation using the mean of reported values. Ideally, in the case of missing data, multiple imputation methods should be employed in order to provide reliable estimates for the missing data values; however, because of the small sample size it is unlikely that complex imputation methods would yield any more reliable results in this instance.

TABLE 58 Proportion of patient-reported questionnaires returned

	Strategy					
	ND			PET–CT surveillance		
Follow-up time point	Received (*n*)	Expected (*n*)	Missing (%)	Received (*n*)	Expected (*n*)	Missing (%)
Within 2 weeks of CRT	39	42	7	38	40	5
3 months post-CRT assessment	37	42	12	38	40	5
6 months post randomisation	36	40	10	36	37	3
12 months post randomisation	36	36	0	37	37	0
24 months post randomisation	34	34	0	35	35	0

Cost-effectiveness analysis

Cost-effectiveness was measured in terms of the ICER and INB. The ICER is calculated by dividing the difference in mean cost between the two arms by the difference in mean QALYs between the two arms:

$$ICER = \frac{C_i - C_c}{E_i - E_c} = \frac{\Delta C}{\Delta E}, \quad (1)$$

where C_i and E_i are the expected cost and effectiveness of the intervention (PET–CT surveillance), C_c and E_c are the expected cost and effectiveness of the comparator strategy (planned ND), and ΔC and ΔE are the incremental cost and effect of the intervention compared with the comparator.

When the new intervention dominates the standard care strategy (i.e. is more effective and less costly), or when the new intervention is dominated by the standard care strategy (i.e. is less effective and more costly), the decision of whether or not to adopt the new strategy is straightforward (accept and reject, respectively). In these cases the ICER calculation is meaningless. The ICER becomes important when we are considering adopting an intervention that requires a trade-off between additional effect for additional cost and less effect for less cost. Assuming that a new intervention is more costly and more effective than the current standard care strategy, the ICER represents the additional cost associated with the intervention per additional unit of benefit (QALY) that it provides compared with current treatment; alternatively, if the intervention is less costly and less effective than the standard care intervention, the ICER represents the cost saved per QALY lost compared with current treatment. The cost-effectiveness of an intervention is then determined by whether or not the ICER value falls above or below the decision-maker's willingness to pay per QALY. In the UK, NICE adopts a willingness-to-pay threshold value of £20,000 per QALY: if a new intervention has an ICER value below £20,000 per additional QALY (or > £20,000 saved per QALY lost), then it is likely to be considered a cost-effective use of NHS resources, whereas an ICER value above £20,000 per additional QALY (or < £20,000 saved per QALY lost) indicates that the intervention is not expected to be a cost-effective use of resources. For the case of positive incremental cost and QALYs (i.e. the intervention is more effective and more costly), this decision rule is expressed in the following formula:

$$\frac{\Delta C}{\Delta E} < \lambda, \quad (2)$$

where λ is the adopted willingness-to-pay threshold (£ per QALY).

When the threshold is known, cost-effectiveness can be expressed in terms of net health benefit (NB). The NB for the given intervention (NB_i) and comparator strategy (NB_c) and the INB of the intervention are calculated as follows:[77]

$$NB_i = E_i - \frac{\Delta C_i}{\lambda}$$

$$NB_c = E_c - \frac{\Delta C_c}{\lambda} \quad (3)$$

$$INB = NB_i - NB_c = \Delta E - \frac{\Delta C}{\lambda}.$$

The NB equation expresses the overall benefit of a strategy in terms of QALYs by converting the expected cost of the intervention on to the QALY scale using the given threshold value. Unlike the ICER, for which interpretation depends on whether or not the incremental cost and QALYs are positive or negative, the interpretation of NB is straightforward: for any given set of strategies, the strategy with the highest net benefit is the most cost-effective; equivalently, a strategy is cost-effective if its INB is positive. All results of the economic evaluation are reported in terms of both ICER and NB values.

Uncertainty and sensitivity analyses

Non-parametric bootstrapping was used to determine the level of sampling uncertainty around the WT cost-effectiveness result by generating 10,000 estimates of incremental costs and benefits from the trial results. The bootstrap approach considers the original sample as if it is the population and draws multiple random samples with replacements from the original sample in order to simulate possible alternative sample sets. Results are presented using scatterplots on the cost-effectiveness plane (which plots incremental QALYs against incremental costs) to illustrate the uncertainty surrounding the cost-effectiveness estimates.

On the cost-effectiveness plane a result is considered cost-effective if it falls on or below the given cost-effectiveness willingness to pay per QALY threshold. The cost-effectiveness acceptability curve is derived by calculating the proportion of bootstrapped estimates that are cost-effective across a range of willingness-to-pay thresholds, to show the probability that PET–CT-guided treatment is cost-effective across different threshold values.

For the WT analysis, three additional sensitivity analyses were conducted to explore additional uncertainties in the analysis.

Within-trial sensitivity analysis 1: imputation of additional patient-reported cost data (NHS and Personal Social Services perspective)

In the base-case analysis, costs were calculated using data from the trial case report forms that provided information on a range of secondary care resource usage for the total trial population ($n = 564$). Additional data on primary and community care resource use, as well as additional secondary care resource usage (outside the enrolled oncology department), were collected for a subgroup of the trial population ($n = 42$) enrolled at the two main recruiting centres (UHB and UHCW). Patients were asked to recall their use of NHS and PSS services over the past 3 months (or since completion of the last form when appropriate) at baseline, during treatment (2 weeks post CRT), and at 3, 6, 12 and 24 months post randomisation. In the base-case analysis this information was excluded because of the small sample size available for these data and the significant uncertainty incorporated in any attempt to impute these data to the whole trial population. A sensitivity analysis was conducted to assess the potential impact of including these additional cost data by imputing the mean reported values for the additional resource use items collected in the patient-reported forms to the total trial population. Mean values were calculated using a complete-case analysis of the patient- and carer-reported data. Owing to the small sample size and significant number of missing data, it was deemed inappropriate to attempt to conduct multiple imputation for missing data.

On the patient-reported questionnaires, patients were asked to report any additional visits to hospital (inpatient, day centre, outpatient, accident and emergency department or nursing home visits), not including visits to their enrolled oncology department; the aim was to capture additional secondary care resource use not already captured in the case report forms routinely completed at the patient's enrolled oncology department. However, when additional hospital visits were reported, patients were asked to give the name of the hospital they visited, and a significant number of patients identified the hospital as UHB or UHCW. The majority of patients did not specify the department visited, so it is unclear if these hospital visits constituted unique events that have not already been captured in the case report forms. Two analyses were therefore conducted:

1. WT sensitivity analysis 1.1: in this analysis it was assumed that patients correctly reported their use of secondary care resources, that is, if the hospital site was reported as UHB or UHCW, patients reported only those hospital events that occurred outside the oncology department and, therefore, each reported event represented a unique resource use item not already contained within the case report forms
2. WT sensitivity analysis 1.2: in this analysis it was assumed that for any instance in which patients identified the hospital visited as either UHB or UHCW these data would already have been captured in the case report forms and these events are therefore excluded from the analysis.

Within-trial sensitivity analysis 2: imputation of additional patient- and carer-reported societal costs (societal perspective analysis)

For the subgroup of patients enrolled at UHB and UHCW, patients and carers were asked to report on the impact of the patient's illness on travel expenses, one-off expenses, equipment costs and lost productivity (i.e. lost earnings), at the same time points as for the patient questionnaires discussed in WT sensitivity analysis 1 (baseline, treatment and 3, 6, 12 and 24 months post randomisation). A sensitivity analysis was conducted to assess the potential impact of including these additional cost data, in addition to the costs included in WT sensitivity analysis 1, by imputing the mean reported values for the additional resource use items collected in the patient- and carer-reported forms to the total trial population. Owing to the fact that the additional costs included in WT sensitivity analysis 1 have been split into two analyses (WT sensitivity analysis 1.1 and sensitivity analysis 1.2), two corresponding analyses have been conducted from a societal perspective:

1. WT sensitivity analysis 2.1, based on inclusion of societal costs plus additional patient-reported costs as given in WT sensitivity analysis 1.1 (i.e. using patient-reported costs as reported)
2. WT sensitivity analysis 2.2, based on inclusion of societal costs plus additional patient-reported costs as given in WT sensitivity analysis 1.2 (i.e. excluding patient-reported secondary care visits where the attended hospital was identified as one of the enrolled hospitals).

Within-trial sensitivity analysis 3: complete-case analysis for calculation of quality-adjusted life-years

In the base-case analysis, multiple imputation methods were used in order to account for missing EQ-5D data within the trial. A sensitivity analysis was conducted using complete-case analysis for the calculation of QALYs in each arm (i.e. ignoring missing data).

Lifetime decision model analysis

In order to estimate the lifetime cost-effectiveness of PET–CT-guided management, a de novo decision-analytical model was constructed. In line with the WT analysis, the base-case model adopts a UK secondary care perspective and future costs and QALYs were discounted at an annual rate of 3.5% in line with NICE guidance.[63] A sensitivity analysis was conducted in which a broader NHS and PSS perspective was adopted. The model was built and analysed using R software.

Model structure

The model structure (*Figure 28*) is split into two phases: an initial 6-month treatment phase, in which patients receive CRT and (potential) ND; and a follow-up phase, in which patients may go on to recover, develop local recurrence (LR) or distant recurrence (DR), or die. Within the treatment period, cost and QALYs for each of the treatment arms are taken directly from the results of the first 6 months of the WT analysis. For the follow-up period, longer-term cost and QALYs for each of the treatment arms are estimated using a modified Markov model. The model begins at this point in order to enable the analysis to capture the full cost and utility decrements associated with LRs and DRs, as these data were not fully captured in the trial.

Markov models describe patient progression over time through a pathway of health states, with movement between the health states being triggered by events such as disease progression or death. Resource use and costs are associated with each health state, and patients accumulate costs and health benefits in each state over monthly cycles. The analysis adopts a lifetime horizon, truncated at 100 years.

During the treatment phase of the model, patients in the standard care arm (arm A) receive planned ND either before or after CRT. Patients in the intervention arm (arm B) receive CRT followed by PET–CT at 9–13 weeks post CRT, which dictates whether or not patients go on to receive ND. In the subsequent follow-up phase of the model, patients enter one of four health states: disease free (DF), LR, DR or death. During any cycle in the follow-up period patients in the DF state may remain DF, develop a LR or DR, or die. Similarly, patients in the LR state may recover, develop a DR or die, and patients in the DR state may remain in DR or die. Entry into the LR state is associated with a one-off treatment cost and utility decrement, after which patients either recover (with costs and utility equal to those in the DF state) or advance to the DR state. Patients who enter the DR state similarly incur a one-off treatment cost and further utility decrement and, subsequently, either remain in the DR state (with an ongoing supportive care cost and utility decrement) or die. Once in the DR state, it is assumed that patients cannot recover and either remain in the DR state or die.

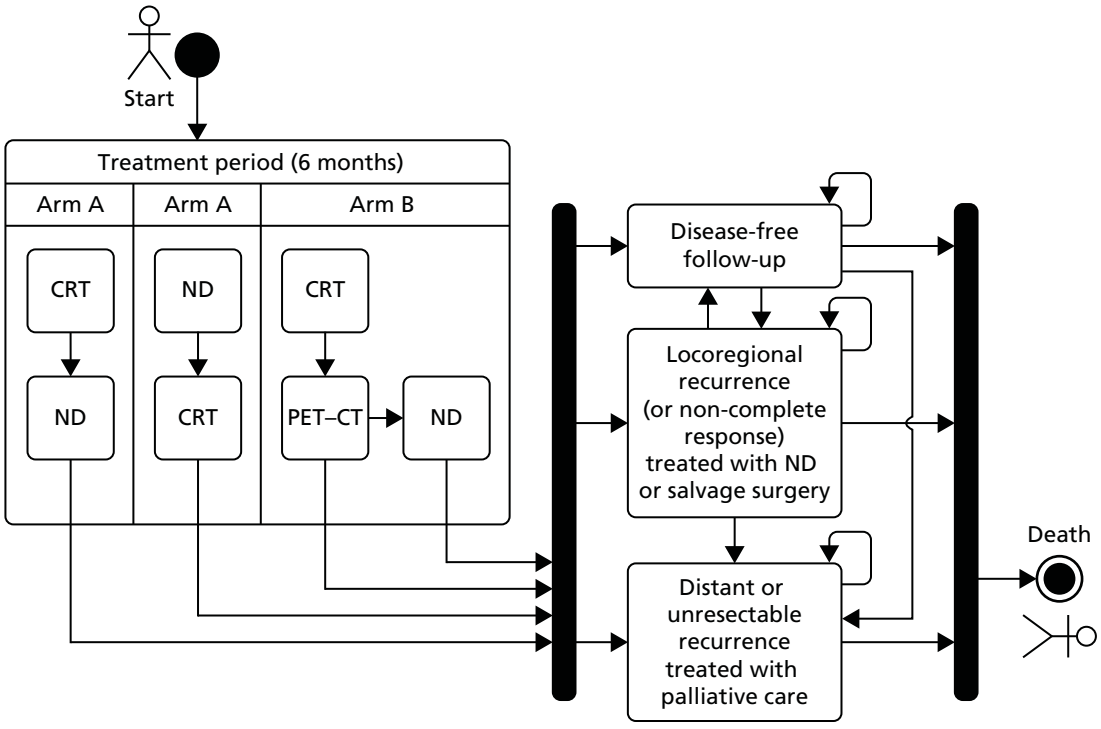


FIGURE 28 Lifetime DM structure.

Model parameters

The model input parameters were derived using information from a range of sources. When possible, data from the trial were used directly and, when necessary, published literature was used to inform remaining parameters.

For the 6-month initial treatment phase of the model, costs and QALYs for each of the arms were taken directly from the trial-reported data. Resource use items relevant to the treatment phase (i.e. surgery, radiotherapy, chemotherapy, treatment response assessment and adverse events and recurrence) were assigned costs and evaluated, as discussed earlier (see *Methods*, *Within-trial analysis*). Bootstrap methods were used to account for sampling uncertainty around the results.

For the follow-up modified Markov phase of the model, the proportion of patients beginning in each of the health states of the model was derived directly from the trial data using data on OS, recurrence-free survival and types of recurrences experienced.

The cost for the DF state was derived from the trial data, based on calculating the average monthly cost in each arm for DF patients over the follow-up period (6–24 months). The cost of LR and DR treatment (applied to patients' first cycle in each of the recurrence states) was derived from the trial data on recurrence treatments administered on patients' first recurrence event. No data were available from the trial regarding the ongoing treatment of patients who experienced a recurrence event; the subsequent cost of LR and DR beyond initial treatment is therefore uncertain. For patients who experienced a LR and, subsequently, recovered after treatment, ongoing costs were assumed to be equal to those for patients in the DF state. For patients remaining in the DR state it was assumed that patients would incur an ongoing supportive care cost, which was derived from the literature.[73]

For the health state utilities, DF utility was derived directly from the trial data by calculating the average utility for patients in the trial who were DF over the second year trial follow-up period (12–24 months). There were no data available from the trial from which the utility of patients experiencing a LR or DR could be derived. Several studies were identified in the literature that have explored health-related quality-of-life for patients with H&N cancers. Most of these studies were deemed inappropriate for the derivation of the current model recurrence health states because the majority either reported only utilities derived from the EORTC QLC-C30 questionnaire (a cancer-specific health-related quality-of-life questionnaire), which cannot be easily converted in to EQ-5D utilities,[78–81] or did not report health state utilities that could be reliably mapped to the health states used in the current model.[82] One UK study[83] was identified. However, this study calculated EQ-5D utilities from a sample of 50 oncology nurses; because health-care professionals are known to be poor proxies for patients, this study was similarly excluded from consideration for the base-case analysis. de Almeida *et al.*[84] conducted an elicitation of utilities from a sample of 50 healthy Canadian subjects using both standard gamble and visual analogue scale (VAS) methods. The authors report utility values for a range of health states, including several remission, LR and DR states. As this study was deemed to be the best-quality study identified in the literature, it was used to derive the recurrence health state values for the current DM. Probabilistic utility decrement values for the local and DR health states were derived by sampling from the remission, LR and DR distributions in the de Almeida *et al.*[84] study and taking the resulting mean differences and standard error values. The base-case analysis uses the reported utilities in de Almeida *et al.*[84] based on the standard gamble elicitation. A sensitivity analysis was conducted using the VAS-reported values.

The mortality risk within the DF, LR and DF after LR states was assumed to be equal to the mortality risk within the general population (taken from Office for National Statistics 2013 statistics[85]) multiplied by an excess mortality factor of 20% derived from the literature.[86] The mortality risk within the DR state was derived by calibrating the model survival curve with the Kaplan–Meier survival data from the trial. The rate of disease progression from the DF state to LR and DR within the first 5 years was derived directly from the trial data, using the trial recurrence-free survival Kaplan–Meier data to derive monthly primary recurrence

transition probabilities, accounting for loss to follow-up. Stochastic primary recurrence probabilities were derived by simulating 10,000 bootstrap data samples from the original Kaplan–Meier data (i.e. sample with replacement), which were used to derive separate recurrence probabilities for each of the 10,000 model simulations in the probabilistic analysis (see *Lifetime decision model analysis*, *Uncertainty and sensitivity analyses*). In the base-case analysis recurrence events were assumed to occur only within the first 5 years post randomisation, based on observations within the literature, which indicate that the majority of recurrences occur within the first 5 years from initial diagnosis.[87] This assumption was also supported by the trial follow-up data, which indicated that the number of recurrence events had fallen close to zero at 5 years post treatment. A sensitivity analysis was conducted to assess the potential impact of recurrences beyond 5 years (see *Lifetime decision model analysis*, *Uncertainty and sensitivity analyses*). The proportion of patients experiencing a LR or DR in each arm was derived from the trial data by calculating the relative proportion of LR versus DR events within the trial. This proportion was assumed to be constant over time in the model.

There were no data available from the trial regarding the rate of progression from LR and DF after LR to subsequent LR or DR (i.e. only primary recurrences were captured in the trial follow-up). The rate of subsequent recurrences were therefore derived from a recent study[88] that reported observed LRs and DRs in 176 patients with local recurrent disease after primary curative treatment of HNSCC.

A full list of model parameters and distributions applied in the base-case model is given in *Table 59*.

Cost-effectiveness analysis

As for the WT analysis, cost-effectiveness was measured in terms of the ICER and INB (see *Lifetime decision model analysis*, *Uncertainty and sensitivity analyses* for more details on the calculation of ICERs and INB).

Uncertainty and sensitivity analyses

Probabilistic sensitivity analyses were conducted to assess the impact of joint parameter uncertainty on the results. Probabilistic analysis accounts for joint parameter uncertainty in non-linear models by assigning probability distributions to each of the input parameters and randomly drawing from these probabilities over 10,000 Monte Carlo model simulations to produce different cost and QALY estimates in each simulation of the model. As for the bootstrap WT analysis, the results are presented on the cost-effectiveness plane as a scatterplot, using cost-effectiveness acceptability curves to show the probability that the two strategies are cost-effective across different willingness-to-pay per QALY thresholds.

TABLE 59 Decision model base-case parameters

Parameter	Mean	Distribution	SD	Source
Global parameters				
Discount rate	0.035	Fixed	–	NICE guidance[63]
Start age (years)	57	Fixed	–	PET-NECK baseline trial data
Percentage male	82%	Fixed	–	
Health state starting distribution (i.e. end of 6-month treatment period) *ND*				
Recurrence	0.06	Beta	0.015	PET-NECK trial data
Proportion of recurrences local (vs. distant)	0.35	Beta	0.069	
Dead	0.03	Beta	0.010	

TABLE 59 Decision model base-case parameters (*continued*)

Parameter	Mean	Distribution	SD	Source
PET–CT				
Recurrence	0.05	Beta	0.013	PET-NECK trial data
Proportion of recurrences local (vs. distant)	0.41	Beta	0.066	
Dead	0.02	Beta	0.008	
Monthly health state costs				
DF	£70	Gamma	£95	PET-NECK trial data on DF patients ($n = 439$)
LR initial treatment	£3964	Gamma	£4343	PET-NECK trial data on LR patients ($n = 39$)
DF after LR	£70	Gamma	£95	Assumed equivalent to DF cost
DR initial treatment	£3635	Gamma	£3197	PET-NECK trial data on DR patients ($n = 63$)
DR follow-up (ongoing care)	£140	Gamma	£32	Hall *et al.*, 2014[73]
Terminal-month cost	£1051	Gamma	£115	Hall *et al.*, 2014[73]
Health state utilities				
DF	0.71	Beta	0.04	PET-NECK trial data on DF patients during second year of follow-up
LR decrement	–0.11	Beta	0.12	de Almeida *et al.*, 2014[84]
DF after LR	0.71	Beta	0.03	Assumed equivalent to DF utility
DR decrement	–0.47	Beta	0.20	de Almeida *et al.*, 2014[84]
Dead	0	Fixed	–	
Transition probabilities/effects				
Primary recurrence over first 5 years	Derived from the trial Kaplan–Meier curves for recurrence-free survival in each arm			PET-NECK trial data
Proportion of recurrences local (vs. distant) in planned ND arm	0.35	Beta	0.069	PET-NECK trial
Proportion of recurrences local (vs. distant) in PET–CT surveillance arm	0.41	Beta	0.066	
Probability of LR from DF after LR state	0.02	Beta	0.002	Matoscevic *et al.*, 2014[88]
Probability of DR from LR/DF after LR states	0.02	Beta	0.003	Matoscevic *et al.*, 2014[88]
Baseline mortality in DF and LR states	Life table	Fixed	–	Office for National Statistics,[85] 2013 age- and sex-standardised rates
Excess mortality factor for DF and LR states	1.20	Fixed	–	van der Schroeff *et al.*, 2010[86]
Mortality in DR state	0.30	Beta	0.30	Calibration of model survival curve against the trial survival data

SD, standard deviation.

One-way sensitivity analyses were conducted to test the impact of changes in individual parameter estimates on the results. Individual parameters were altered by ± 25% from their baseline value, with the impact on the expected INB reported in a tornado diagram.

For the lifetime DM analysis, three additional sensitivity analyses were conducted to assess the impact of key assumptions used in the model.

Decision model sensitivity analysis 1: imputation of additional patient-reported cost data (NHS and personal social services perspective)

As for the WT cost-effectiveness analysis, the base-case DM analysis was based on utilisation of data from the trial relating to patients' consumption of NHS secondary care resources, using the trial case report forms, which were collected for the total trial population. In line with the sensitivity analyses conducted in the WT sensitivity analysis 1.1 and sensitivity analysis 1.2 (see *Within-trial analysis*, *Uncertainty and sensitivity analyses*), two corresponding sensitivity analyses were conducted in the model analysis to assess the impact of including additional patient-reported cost data:

1. DM sensitivity analysis 1.1: in this analysis the treatment-phase costs for the model were taken from the treatment-phase costs generated from the WT sensitivity analysis that included the additional patient-reported cost data, assuming that patients correctly reported their use of additional secondary care resources (i.e. corresponding to analysis WT sensitivity analysis 1.1). Similarly, the monthly cost of the DF state was recalculated using the results of WT sensitivity analysis 1.1. In addition, the average monthly patient-reported cost associated with the use of primary, community and additional secondary care (i.e. outside the enrolled oncology department) was applied to the local and DR health states. As this is likely to underestimate the primary and community care costs for these states (because the added cost data are derived from primarily DF patients), no other data were available to determine the exact cost of primary and community care for recurrent patients.
2. DM sensitivity analysis 1.2: this analysis involved the same steps as outlined in DM sensitivity analysis 1.1 above, but instead used data from the WT sensitivity analysis, including additional patient-reported data and excluding secondary care resource use where patients identified the visited hospital as one of the trial centres (i.e. WT sensitivity analysis 1.2; see *Within-trial analysis*, *Uncertainty and sensitivity analyses* for further details).

Decision model sensitivity analysis 2: visual analogue scale utility values for the local recurrence and distant recurrence states

In the base-case analysis, utility in the recurrence health states was derived using data reported in de Almeida *et al.*,[84] based on an elicitation of utility values using the standard gamble method. The authors also reported utilities based on using the VAS tool. A sensitivity analysis was conducted using these alternative values in order to assess the impact of the method of utility derivation on the results.

Decision model sensitivity analysis 3: recurrences beyond 5 years post diagnosis estimated using a parametric survival curve

In the base-case analysis, no recurrence events were assumed to occur beyond 5 years in the model. A sensitivity analysis was conducted assuming that recurrences could occur beyond 5 years: long-term recurrence probabilities for the planned ND arm were estimated using a Gompertz parametric survival curve fitted to the trial baseline Kaplan–Meier data on recurrence-free survival; a HR was then applied to this curve in order to derive DF survival within the PET–CT surveillance arm (using the HR observed across the trial follow-up period), as outlined in Briggs *et al.*[89] The Gompertz distribution was identified as the best-fitting curve to estimate long-term recurrence events (compared with the exponential, Weibull, gamma and log-normal distributions), based on an analysis of the Akaike information criterion (AIC) and Bayesian information criterion (BIC) (*Table 60*; better-fitting curves are indicated by lower AIC and BIC values) and a visual inspection of the curve fits (*Figure 29*).

TABLE 60 Akaike information criterion and BIC to estimate the goodness of fit of alternative parametric survival curve model specifications for the estimation of long-term recurrences

	Model specification	
Trial arm	AIC	BIC
Planned ND		
Exponential	679.50	683.14
Gamma	674.96	682.96
Weibull	673.83	681.11
Log-normal	665.86	673.14
Gompertz	659.03	666.31
PET–CT surveillance		
Exponential	734.08	737.72
Gamma	735.58	742.86
Weibull	735.16	742.44
Gompertz	727.89	735.17
Log-normal	726.19	733.48

Results

Within-trial analysis

Within-trial costs and utilities

Costs: base case (NHS secondary care perspective)

A summary of the 2-year WT costs included in the base-case analysis (i.e. NHS secondary care costs) is provided in *Table 61*.

The total per-patient costs over the 2-year trial follow-up period were significantly higher in the standard care (planned ND) arm than in the PET–CT surveillance arm (mean £13,989 vs. £12,476; $p = 0.000$). This is primarily because of the higher up-front cost of surgery (£2542 vs. £303) as well as higher adverse event costs (£1167 vs. £655), which result in higher treatment-phase costs for planned ND than for PET–CT surveillance (£12,289 vs. £10,578). In contrast, subsequent costs in the follow-up period (6–24 months post randomisation) were non-significantly higher in the PET–CT arm (mean £1898 vs. £1700, $p = 0.24$). This is mainly because of higher costs associated with surgery (£467 vs. £266) and recurrence (£671 vs. £496) in the PET–CT arm during the follow-up phase.

Costs: sensitivity analysis 1 (NHS and Personal Social Services perspective)

A summary of the patient-reported costs used in sensitivity analysis 1 is presented in *Table 62*. These costs include patient-reported primary, (additional) secondary and community care costs collected for a subset ($n = 42$) of patients enrolled at the two main recruiting centres (UHB and UHCW). WT sensitivity analysis 1.1 includes patient-reported costs as they were reported; WT sensitivity analysis 1.2 excludes hospital visits for patients who identified the visited hospital as either UHB or UHCW (because of the likelihood that these events will be duplicated in the trial case report forms).

For both analyses the total mean patient-reported costs were higher in the planned ND arm; however, the cost difference was greater in WT sensitivity analysis 1.1 (£18,245 vs. £13,869) than in WT sensitivity

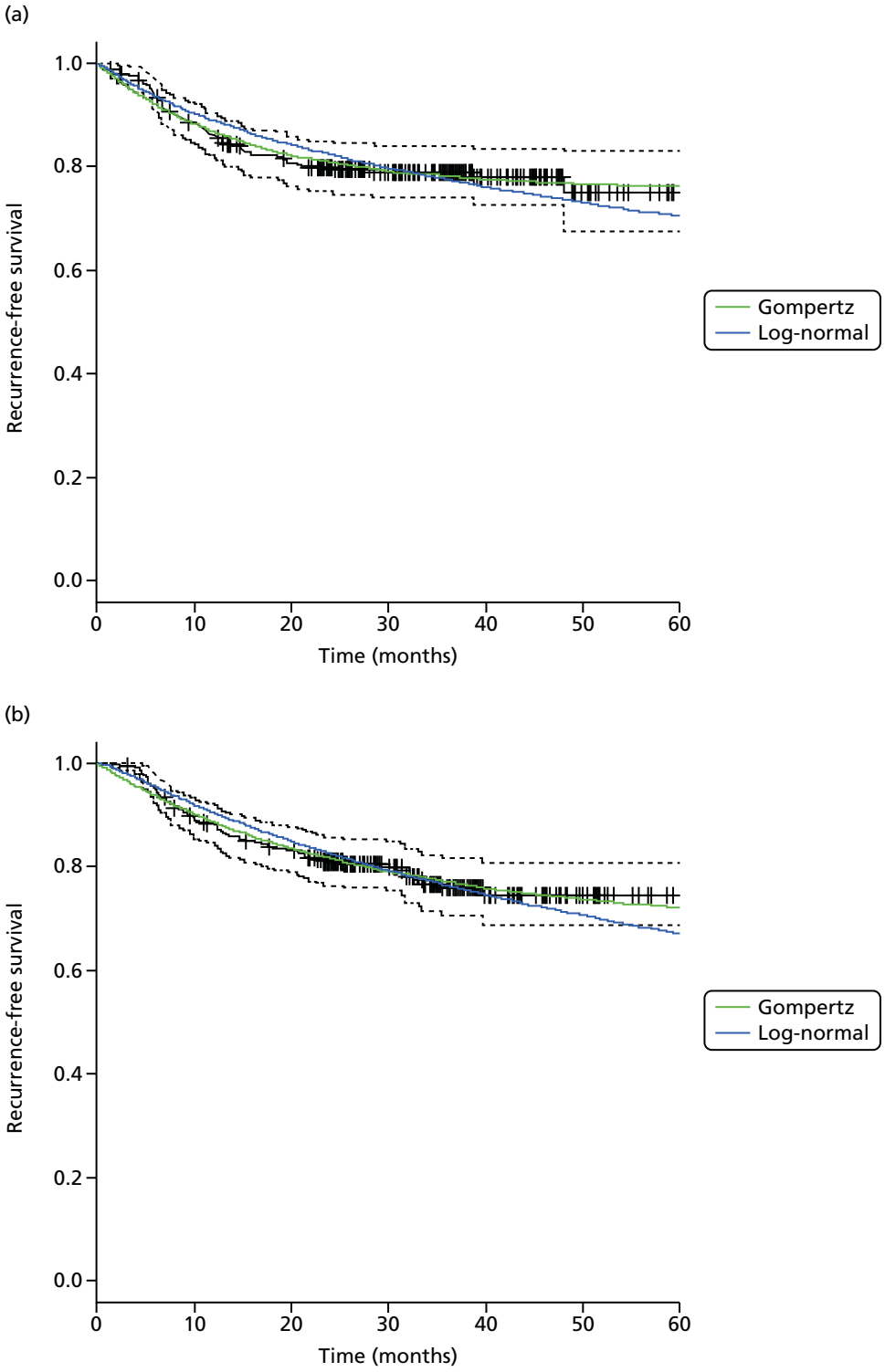


FIGURE 29 Plot of parametric survival curves (using Gompertz and log-normal model specifications) against trial data on patient recurrence-free survival. (a) Planned ND arm; and (b) PET–CT surveillance arm. Log-normal and Gompertz specifications are shown here, as these were the two models identified as having the best fit via an analysis of AIC and BIC. The Gompertz curve was used in the sensitivity analysis, as this had the best overall fit when taking into account AIC, BIC and a visual inspection of the goodness of fit.

TABLE 61 Within-trial base-case costs: summary of the average per patient secondary care costs by treatment phase

Cost category	Strategy (£)						Mean difference (£)	*p*-value (Wilcoxon signed-rank test)
	Planned ND ($n = 282$)			PET–CT ($n = 282$)				
	Mean	Median	SD	Mean	Median	SD		
Treatment period (0–6 months)								
Surgery	2542	3566	1616	303	0	997	–2238	0.000
Radiotherapy	5768	6087	1677	5963	6087	1334	194	0.451
Chemotherapy: concomitant	1444	724	1811	1527	764	1837	82	0.361
Chemotherapy: neoadjuvant	774	0	1252	922	0	1313	148	0.224
Treatment assessment	437	508	183	354	360	123	–83	0.000
PET–CT	69	0	230	833	652	358	763	0.000
SAE	1167	0	2081	655	0	1185	–512	0.010
Recurrence	88	0	633	21	0	151	–67	0.166
Total treatment period cost	12,289	12,052	3969	10,578	10,081	3155	–1711	0.000
Follow-up period (6–24 months)								
Surgery	266	0	938	467	0	1203	201	0.029
Radiotherapy	24	0	407	0	0	0	–24	0.319
Follow-up assessments	416	484	223	449	484	192	33	0.271
PET–CT	70	0	237	103	0	274	34	0.065
SAE	429	0	1316	208	0	706	–221	0.146
Recurrence	496	0	1846	671	0	2181	175	0.248
Total follow-up costs	1700	542	5041	1898	589	2970	198	0.243
Total costs (0–24 months)								
Surgery	2807	3566	1462	770	0	1468	–2037	0.000
Radiotherapy	5793	6087	1643	5963	6087	1334	170	0.533
Chemotherapy: concomitant	1444	724	1811	1527	764	1837	82	0.361
Chemotherapy: neoadjuvant	774	0	1252	922	0	1313	148	0.224
Treatment assessment	437	508	183	354	360	123	–83	0.000
Follow-up assessments	416	484	223	629	652	121	214	0.271
PET–CT	139	338	0	936	652	464	797	0.000
SAE	1596	0	3057	863	0	1690	–733	0.609
Recurrence	584	0	1932	693	0	2181	109	0.609
Total overall cost	13,989	13,278	5041	12,476	11,940	4613	–1513	0.000

SD, standard deviation.

TABLE 62 Within-trial patient-reported costs included in sensitivity analysis WT sensitivity analysis 1.1 and WT sensitivity analysis 1.2

Questionnaire time point	Strategy (£)								Mean difference (£)	*p*-value (Wilcoxon signed-rank test)
	Planned ND				PET–CT surveillance					
	n	Mean	Median	SD	*n*	Mean	Median	SD		
Sensitivity analysis 1.1: analysis using patient data as reported										
Treatment period	17	3170	3261	2059	25	2244	1191	1994	–926	0.118
3 months post randomisation	15	3225	3183	1898	20	1098	1007	501	–2127	0.000
6 months post randomisation	4	2009	2166	1281	10	1076	928	814	–933	0.240
12 months post randomisation	16	1224	825	1406	21	1125	872	1137	–99	0.963
24 months post randomisation	15	1183	797	1406	14	1087	843	1099	–96	0.963
Total[a]	–	18,245	18,245	1469	–	13,869	13,869	1494	–4377	0.000
Sensitivity analysis 1.2: analysis excluding patient-reported secondary care resource usage in which the visited hospital was indicated as UHB or UHCW										
Treatment period	17	962	753	667	25	1032	846	720	70	0.547
3 months post randomisation	15	1216	787	914	20	883	787	304	–333	0.802
6 months post randomisation	4	855	843	502	10	748	751	358	–107	0.733
12 months post randomisation	16	1121	721	1430	21	935	707	684	–186	0.914
24 months post randomisation	15	1083	696	667	14	904	683	661	–179	0.929
Total[a]	–	11,291	11,219	1355	–	10,193	10,193	810	–1098	0.000

SD, standard deviation.

a Total cost after imputation of mean costs to the total trial population. Calculation of total cost required extrapolation between time points as questionnaires were limited to 3-month recall periods. The total assumes that the 9-month cost is equal to the mean of the 6- and 12-month costs; the 15-month cost is equal to the 12-month cost; the 18-month cost is equal to the mean of the 12-month and 24-month costs; and the 21-month cost is equal to the 24-month cost.

analysis 1.2 (£11,291 vs. £10,193). Note that because of the small numbers of completed questionnaires at each time point, particularly within the planned ND arm, these results are subject to significant sampling uncertainty (as is reflected in the large standard deviation values) and should therefore be interpreted with caution.

Costs: sensitivity analysis 2 (societal perspective)

A summary of the patient- and carer-reported societal costs used in sensitivity analysis 2 is presented in *Table 63*. These costs include reported equipment costs, one-off expenses, transport costs and lost earnings. As for the patient-reported costs reported in *Table 62*, the reported societal costs were higher in the planned ND arm than in the PET–CT surveillance arm, and this difference was most substantial for the patient-reported costs (total patient costs £17,136 vs. £11,567; total carer costs £6762 vs. £5388). However, as for the previous sensitivity analysis, these costs should be interpreted with caution given the substantial uncertainty because of the small sample sizes contributing to each of the cost estimates.

TABLE 63 Within-trial patient- and carer-reported societal costs included in WT sensitivity analysis 2

Questionnaire time point	Strategy (£)								Mean difference	*p*-value (Wilcoxon signed-rank test)
	Planned ND				PET–CT surveillance					
	n	Mean	Median	SD	*n*	Mean	Median	SD		
Sensitivity analysis 2: patient costs										
Treatment period	17	1951	1211	3026	25	1547	34	2530	–404	0.421
3 months post randomisation	15	3462	58	6756	20	2804	20	5193	–657	0.891
6 months post randomisation	3	3475	1200	5468	10	940	103	1394	–2535	0.942
12 months post randomisation	16	1932	482	2979	21	1831	1	3170	–101	0.747
24 months post randomisation	15	286	0	871	14	125	0	458	–161	0.649
Total[a]	–	17,136	17,136	4018	–	11,567	11,567	4013	–5569	0.000
Sensitivity analysis 2: carer costs										
Treatment period	13	1644	1688	859	22	1732	1365	1159	88	0.733
3 months post randomisation	10	445	414	135	16	849	528	760	405	0.861
6 months post randomisation	3	1538	1723	395	8	416	209	527	–1122	0.085
12 months post randomisation	10	381	302	469	14	451	253	516	70	0.861
24 months post randomisation	6	294	212	284	11	276	119	347	–18	0.733
Total[a]	–	6762	6762	429	–	5388	5388	660	–1374	0.000

SD, standard deviation.

a Total costs calculated after imputation of sensitivity analysis mean costs to total trial population. Calculation of total cost required extrapolation between time points as questionnaires were limited to 3-month recall periods. The total assumes that the 9-month cost is equal to the mean of the 6- and 12-month costs; the 15-month cost is equal to the 12-month cost; the 18-month cost is equal to the mean of the 12-month and 24-month costs; and the 21-month cost is equal to the 24-month cost.

Utilities

A summary of the mean utilities in each arm over time based on both the base-case analysis (multiple imputation for missing data) and sensitivity analysis 3 (complete-case analysis) is given in *Table 64*. Patient-reported utilities were largely similar between the two arms, with the largest difference occurring at 3 months post randomisation.

Base-case cost-effectiveness results

The results of the WT cost-effectiveness analysis are presented in *Table 65* (deterministic results) and *Table 66* (probabilistic results). Over the 2-year trial period, PET–CT-guided management was associated with an expected incremental cost saving of £1492 and an expected incremental effect of +0.08 QALYs compared with planned ND. PET–CT surveillance was therefore found to be cost-effective and dominates the standard care strategy, being less costly and more effective with an INB of +0.16 QALYs (95% CI 0.03 to 0.28 QALYs).

TABLE 64 Patient-reported utilities over time (non-discounted) for imputed utilities (used in the WT base-case analysis) and complete case utilities (used in sensitivity analysis 3)

Time point	Statistic	Strategy (sensitivity analysis 3)		*p*-value (base case; Wilcoxon signed-rank test)
		Planned ND	PET–CT	
Baseline	Mean	0.76 (0.76)	0.76 (0.77)	0.884
	SD	0.24 (0.23)	0.24 (0.23)	
	Median	0.80 (0.80)	0.80 (0.80)	
Treatment	Mean	0.55 (0.50)	0.49 (0.52)	0.764
	SD	0.27 (0.30)	0.31 (0.30)	
	Median	0.62 (0.62)	0.62 (0.62)	
3 months post randomisation	Mean	0.35 (0.41)	0.67 (0.64)	0.000
	SD	0.39 (0.38)	0.31 (0.32)	
	Median	0.19 (0.62)	0.73 (0.73)	
6 months post randomisation	Mean	0.65 (0.66)	0.68 (0.64)	0.015
	SD	0.28 (0.28)	0.30 (0.32)	
	Median	0.73 (0.73)	0.76 (0.73)	
12 months post randomisation	Mean	0.72 (0.70)	0.70 (0.73)	0.556
	SD	0.30 (0.31)	0.34 (0.31)	
	Median	0.76 (0.75)	0.80 (0.80)	
24 months post randomisation	Mean	0.71 (0.78)	0.74 (0.77)	0.007
	SD	0.29 (0.25)	0.34 (0.31)	
	Median	0.76 (0.80)	0.85 (0.85)	

SD, standard deviation.

TABLE 65 Within-trial analysis base-case deterministic cost-effectiveness results (NHS secondary care perspective and imputed QALYs)

Strategy	Total cost (£)	Total QALY	Incremental cost (£)	Incremental QALY	ICER	NB (QALYs)	Incremental NB (QALYs)
Planned ND	13,989	1.20	–	–	–	0.50	–
PET–CT	12,476	1.27	–1513	0.07	Dominant	0.66	0.15

Figure 30 shows the results of the WT probabilistic cost-effectiveness analysis on the incremental cost-effectiveness plane. Each of the points represents one of the 10,000 simulated bootstrap results, and provides an indication of the expected uncertainty around the mean cost-effectiveness result. From this graph we can see that in almost every simulation PET–CT surveillance is expected to be cost-saving compared with planned ND (i.e. the majority of points lie below the zero line for the incremental cost results). The expected incremental benefit of PET–CT-guided treatment is more uncertain, with points lying across both the positive and negative incremental benefit domains. At a £20,000 per QALY threshold, PET–CT surveillance is cost saving 99% of the time, and more effective than planned ND 91% of the time.

TABLE 66 Within-trial analysis base-case probabilistic cost-effectiveness results (NHS secondary care perspective, imputed QALYs and 10,000 bootstrap simulations)

Strategy	Total cost (£) (95% CI)	Total QALY (95% CI)	Incremental cost (£) (95% CI)	Incremental QALY (95% CI)	ICER	NB (95% CI)	Incremental NB (95% CI)
Planned ND	13,944 (13,269 to 14,509)	1.19 (1.10 to 1.27)	–	–	–	0.49 (0.41 to 0.58)	–
PET–CT	12,457 (11,964 to 12,961)	1.27 (1.18 to 1.34)	–1492 (–698 to –2239)	0.08 (–0.03 to 0.19)	Dominant	0.65 (0.56 to 0.72)	0.16 (0.03 to 0.28)

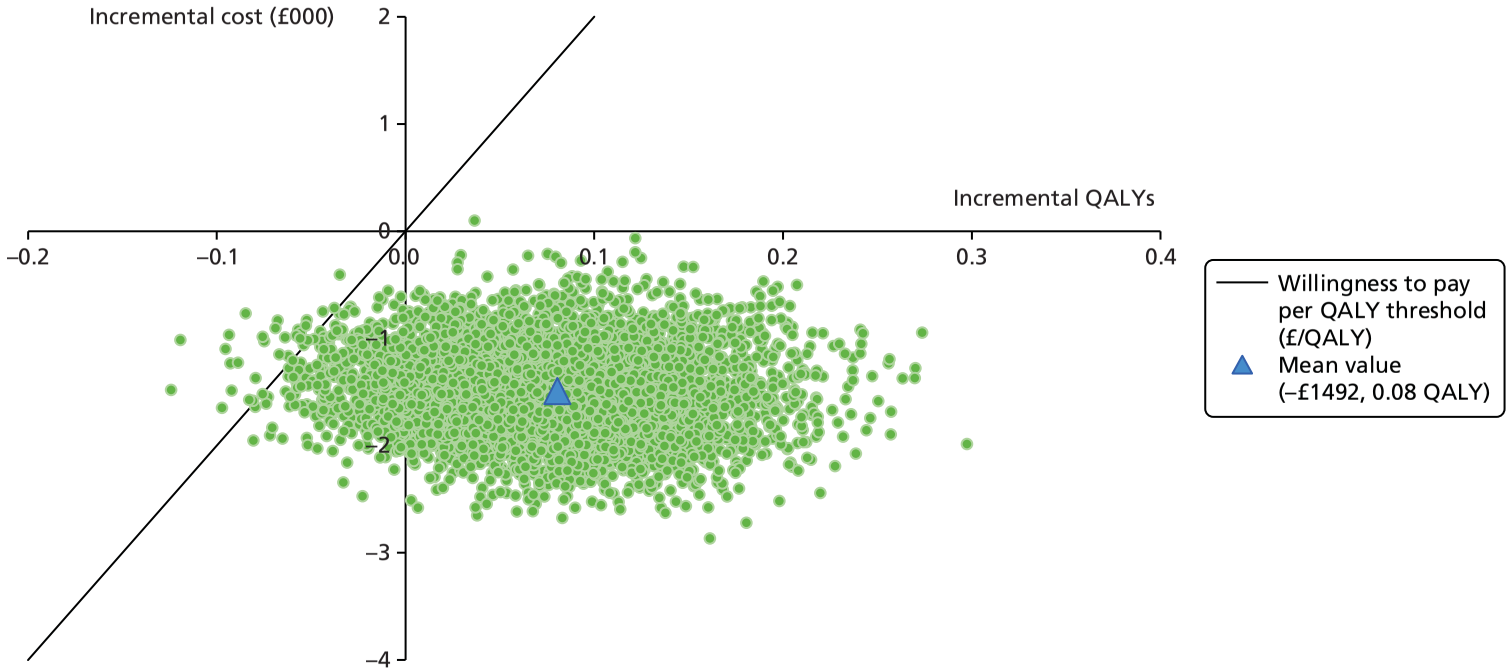


FIGURE 30 Within-trial analysis base-case probabilistic results: incremental cost and QALYs of PET–CT-guided care vs. planned ND (bootstrap analysis and 10,000 simulations).

Figure 31 shows the cost-effectiveness acceptability curve for the WT base-case analysis. This graph shows the probability that each strategy is the most cost-effective alternative (i.e. has the highest expected net benefit) across a range of willingness to pay per QALY thresholds. At a £20,000 per QALY threshold, PET–CT surveillance is associated with a 99% probability of being cost-effective compared with planned ND. This probability remains above 93% up to a £150,000 per QALY threshold.

Sensitivity analysis

Sensitivity analysis 1: imputation of additional patient-reported cost data (NHS and personal social services perspective)

The results of the sensitivity analysis, including the additional patient-reported cost data collected for a subgroup of patients in the trial [including data on resource usage of primary, (additional) secondary and community care services], are presented in *Table 67*. Inclusion of the patient-reported additional costs leads to substantial increases in the expected per patient lifetime cost in both arms, with an increase from £13,944 in the base case to £32,469 in the planned ND arm, and from £12,457 to £26,350 per patient in the PET–CT surveillance arm. As the increase in costs is greater in the planned ND arm than in the PET–CT surveillance arm, the relative cost-effectiveness of PET–CT surveillance is increased, with a rise in the expected INB from 0.16 QALYs to 0.39 QALYs. The probability that PET–CT is cost-effective compared with planned ND is 100% at a £20,000 per QALY thresholds, and remains above 99% up to a threshold value of £150,000/QALY (results not shown).

Table 68 shows the results of the same analysis but with the additional cost data amended to account for potential error in patient-reported values. Again, the expected lifetime cost associated with both arms is increased from the base-case analysis; however, the increase is smaller in this instance and the relative difference in cost increase between arms is reduced. The result is that the INB associated with PET–CT surveillance is slightly reduced to 0.21 QALYs, compared with 0.39 QALYs in the previous analysis.

Within-trial sensitivity analysis 2: inclusion of societal costs (societal perspective)

The results of the WT sensitivity analysis including both the additional costs included in WT sensitivity analysis 1 and patient and carer-reported societal costs (such as travel, equipment and lost earning expenses) are shown in *Tables 69* and *70*.

As in WT sensitivity analysis 1, including societal costs in the analysis leads to substantial increases in the expected cost of each strategy. Again, this cost increase appears to be more pronounced in the planned ND arm, leading to PET–CT surveillance being more cost-effective with increased INB values (to +0.70 QALYs in WT sensitivity analysis 2.1 and to +0.53 QALYs in WT sensitivity analysis 2.2).

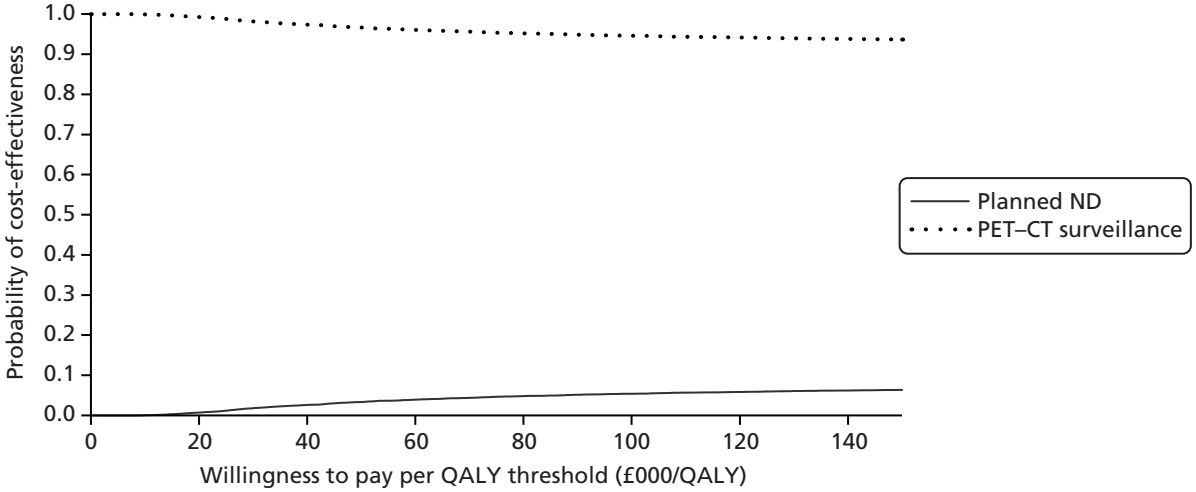


FIGURE 31 Cost-effectiveness acceptability curve for the base-case WT cost-effectiveness analysis.

TABLE 67 Within-trial sensitivity analysis 1.1: WT cost-effectiveness sensitivity analysis using additional cost data (NHS and PSS perspective, additional data used as reported)

Strategy	Total cost (£) (95% CI)	Total QALY (95% CI)	Incremental cost (£) (95% CI)	Incremental QALY (95% CI)	ICER	NB (QALYs) (95% CI)	Incremental NB (QALYs) (95% CI)
Planned ND	32,469 (31,948 to 32,966)	1.18 (1.09 to 1.27)	–	–	–	–0.446 (–0.47 to –0.40)	–
PET–CT	26,350 (25,452 to 26,586)	1.26 (1.19 to 1.33)	–6119 (–6915 to –5270)	0.09 (–0.02 to 0.17)	Dominant	–0.054 (–0.14 to 0.02)	0.39 (0.29 to 0.48)

TABLE 68 Within-trial sensitivity analysis 1.2: WT cost-effectiveness sensitivity analysis using additional cost data (NHS and PSS perspective, excluding secondary care resource use items identifying UHB or UHCW as the visited hospital)

Strategy	Total cost (£) (95% CI)	Total QALY (95% CI)	Incremental cost (£) (95% CI)	Incremental QALY (95% CI)	ICER	NB (QALYs) (95% CI)	Incremental NB (QALYs) (95% CI)
Planned ND	25,274 (24,646 to 25,879)	1.19 (1.09 to 1.27)	–	–	–	–0.08 (–0.17 to 0.01)	–
PET–CT	22,630 (22,134 to 23,186)	1.26 (1.18 to 1.35)	–2645 (–3467 to –1805)	0.08 (–0.05 to 0.21)	Dominant	0.13 (0.03 to 0.24)	0.21 (0.08 to 0.35)

TABLE 69 Within-trial sensitivity analysis 2.1: WT cost-effectiveness sensitivity analysis including patient and carer costs (societal perspective, based on patient-reported costs from WT sensitivity analysis 1.1)

Strategy	Total cost (£) (95% CI)	Total QALY (95% CI)	Incremental cost (£) (95% CI)	Incremental QALY (95% CI)	ICER	NB (QALYs) (95% CI)	Incremental NB (QALYs) (95% CI)
Planned ND	48,753 (48,107 to 49,485)	1.19 (1.09 to 1.27)	–	–	–	–1.25 (–1.35 to –1.16)	–
PET–CT	36,464 (35,780 to 37,126)	1.27 (1.19 to 1.34)	–12,289 (–13,218 to –11,394)	0.08 (–0.03 to 0.19)	Dominant	–0.56 (–0.64 to –0.47)	0.70 (0.57 to 0.82)

TABLE 70 Within-trial sensitivity analysis 2.2: WT cost-effectiveness sensitivity analysis including patient and carer costs (societal perspective, based on patient-reported costs from WT sensitivity analysis 1.2)

Strategy	Total cost (£) (95% CI)	Total QALY (95% CI)	Incremental cost (£) (95% CI)	Incremental QALY (95% CI)	ICER	NB (QALYs) (95% CI)	Incremental NB (QALYs) (95% CI)
Planned ND	41,786 (41,107 to 42,489)	1.19 (1.10 to 1.27)	–	–	–	–0.90 (–0.99 to –0.81)	–
PET–CT	32,799 (32,175 to 33,472)	1.27 (1.19 to 1.34)	–8987 (–9928 to –8047)	0.08 (–0.04 to 0.19)	Dominant	–0.38 (0.46 to –0.28)	0.53 (0.39 to 0.65)

Within-trial sensitivity analysis 3: complete-case analysis for quality-adjusted life-years calculation

The results of the sensitivity analysis using complete-case analysis for the WT QALY calculation are reported in *Table 71*. The PET–CT-guided watch-and-wait policy remains the dominant strategy, being more effective and less costly than planned ND. The expected incremental cost and QALY outcomes are very similar to the base case and result in the same expected INB (0.16 QALYs).

The results of this analysis indicate that the use of multiple imputation versus complete-case analysis for the calculation of QALYs in the WT analysis has minimal impact on the overall results. As a result of this finding, no corresponding sensitivity analysis using complete-case QALYs in the DM was conducted.

Lifetime decision model analysis

Base-case results

Results of the lifetime DM base-case analysis are presented in *Table 72*. Compared with planned ND, PET–CT surveillance is expected to lead to a lifetime cost saving of –£1485 and a health gain of +0.13 QALYs per patient. The intervention therefore dominates standard care, being more effective and less costly than planned ND and resulting in an expected INB of 0.21 QALYs. There is, however, uncertainty around this result, with a wide CI around the mean value (INB 95% CI –0.41 to +0.85 QALYs).

Figure 32 shows the results of the model base-case analysis on the incremental cost-effectiveness plane. Each of the points represents one of the 10,000 simulated model results and provides an indication of the expected uncertainty around the mean cost-effectiveness result. The diagonal line represents NICE's willingness to pay per QALY threshold of £20,000 per QALY: points lying under the line indicate simulations where the intervention is expected to be cost-effective, whereas points lying above the line are not cost-effective. This graph illustrates the increased uncertainty around the lifetime expected incremental cost and QALY results compared with the WT analysis, with points being widely distributed across both cost-effective and non-cost-effective areas of the graph. The majority of the uncertainty lies around the expected effectiveness of the PET–CT surveillance strategy: at a willingness to pay per QALY threshold of £20,000 per QALY, the probability that PET–CT is cost saving compared with planned ND is 96% (compared with 99% in the WT analysis), whereas the probability that PET–CT is more effective than planned ND is 66% (compared with 91% in the WT analysis).

The uncertainty around the lifetime modelled results is reflected in the cost-effectiveness acceptability curve shown in *Figure 33*.

This graph shows the probability that each strategy is the most cost-effective alternative (i.e. has the highest expected net benefit) across a range of willingness to pay per QALY thresholds. At a £20,000 per QALY threshold, PET–CT surveillance is associated with a 75% probability of being cost-effective compared with planned ND (compared with 99% in the WT analysis), dropping to 68% at a £100,000/QALY threshold, and remaining above 67% at a £150,000/QALY threshold.

Sensitivity analysis

One-way sensitivity analysis: tornado plot

The results of the one-way sensitivity analysis, altering individual model parameters by 25%, are presented in *Figure 34*. The results are most sensitive to changes in the treatment period costs and QALYs, and parameters concerned with the rate of recurrences in each arm. In particular, increasing the rate of primary recurrences in the PET–CT surveillance arm (or decreasing the rate of primary recurrences in the planned ND arm) results in the PET–CT watch-and-wait strategy no longer being cost-effective, with the INB falling below zero. Changes to other cost and utility parameters had little impact on the results.

TABLE 71 Within-trial sensitivity analysis 2: WT cost-effectiveness sensitivity analysis using complete-case analysis for QALY calculation

Strategy	Total cost (£) (95% CI)	Total QALY (95% CI)	Incremental cost (£) (95% CI)	Incremental QALY (95% CI)	ICER	NB (QALYs) (95% CI)	Incremental NB (QALYs) (95% CI)
Planned ND	13,985 (13,395 to 14,584)	1.22 (1.15 to 1.29)	–	–	–	0.52 (0.44 to 0.60)	–
PET–CT	12,481 (11,946 to 13,031)	1.31 (1.24 to 1.37)	–1504 (–2312 to –709)	0.09 (–0.01 to 0.19)	Dominant	0.68 (0.61 to 0.76)	0.16 (0.06 to 0.27)

TABLE 72 Lifetime DM base-case analysis (NHS secondary care perspective)

Strategy	Total cost (£) (95% CI)	Total QALY (95% CI)	Incremental cost (£) (95% CI)	Incremental QALY (95% CI)	ICER	NB (QALYs) (95% CI)	Incremental NB (QALYs) (95% CI)
Planned ND	24,074 (12,947 to 63,200)	9.01 (7.87 to 10.46)	–	–	–	7.80 (5.65 to 9.32)	–
PET–CT	22,589 (11,319 to 62,155)	9.14 (8.05 to 10.55)	–1485 (–2815 to 159)	0.13 (–0.49 to 0.79)	Dominant	8.01 (5.85 to 9.51)	0.21 (–0.41 to 0.85)

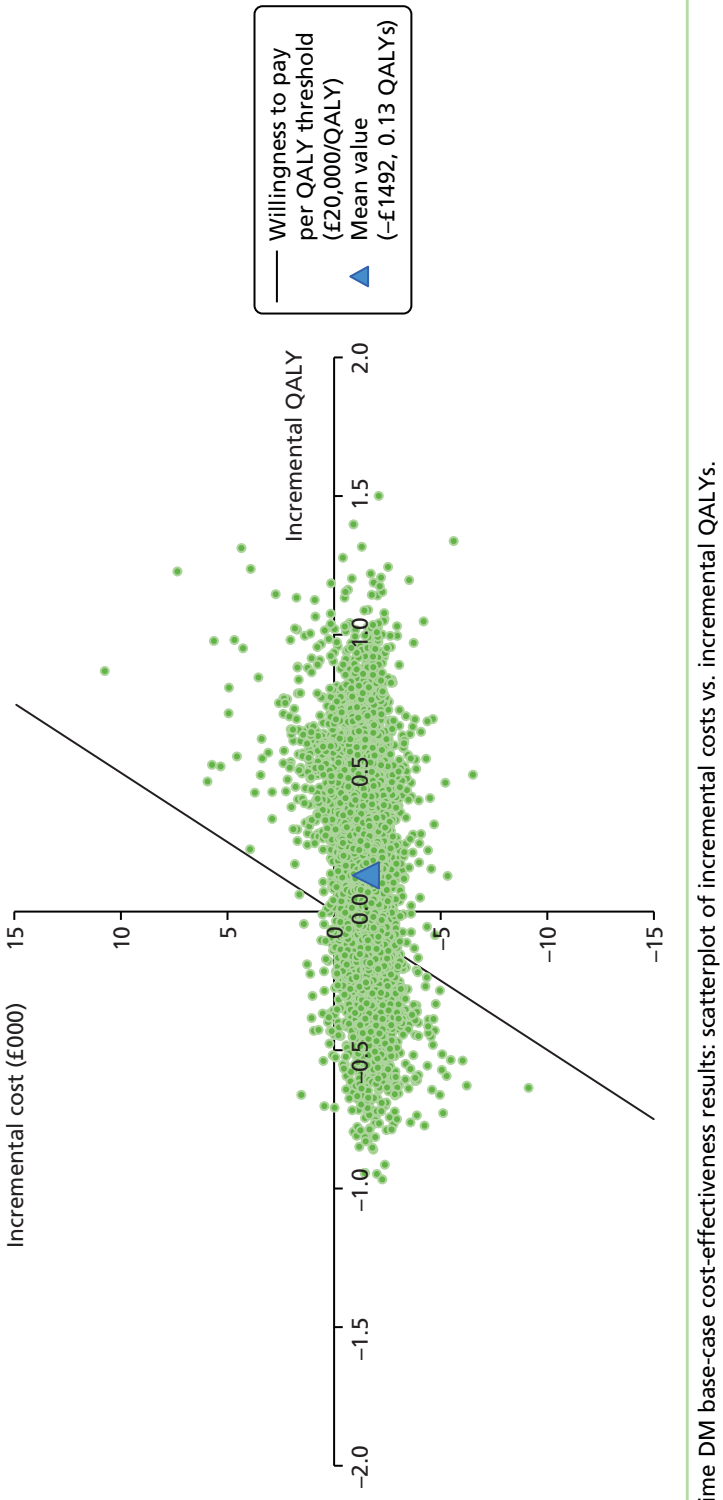


FIGURE 32 Lifetime DM base-case cost-effectiveness results: scatterplot of incremental costs vs. incremental QALYs.

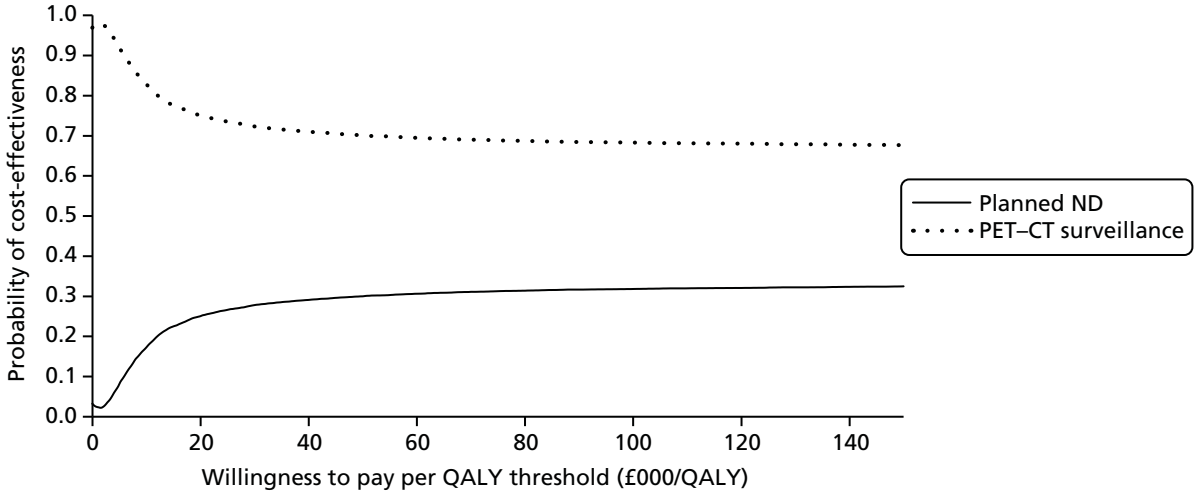


FIGURE 33 Lifetime DM: cost-effectiveness acceptability curve.

Decision model sensitivity analysis 1: inclusion of additional patient-reported cost data (NHS and personal social services perspective)

The results of DM sensitivity analysis 1 are presented in *Table 73* (DM sensitivity analysis 1.1) and *Table 74* (DM sensitivity analysis 1.2). As for the base-case analysis, PET–CT surveillance dominates standard care (i.e. is cost saving and more effective); however, the expected costs of each strategy are substantially higher than in the base-case analysis. In DM sensitivity analysis 1.1, the per patient lifetime expected costs of planned ND and PET–CT surveillance increase from £24,074 and £22,589 to £125,147 and £120,872, respectively, and in DM sensitivity analysis 1.2 the costs increase to £99,898 and £99,198, respectively. At a £20,000 per QALY threshold the probability that PET–CT surveillance is the most cost-effective strategy is 98% in DM sensitivity analysis 1.1 and 81% in DM sensitivity analysis 1.2 (compared with 75% in the base case). In DM sensitivity analysis 1.1, this value remains above 70% up to a threshold of £150,000/QALY, whereas in DM sensitivity analysis 1.2, this value remains above 67%.

Although these results suggest that PET–CT surveillance remains cost-effective over a broader NHS and PSS perspective, the results should be considered as exploratory only because of the substantial uncertainty around the additional costs used to inform this analysis (which were based on data from a small subpopulation of the trial cohort).

Decision model sensitivity analysis 2: visual analogue scale utility values for the local recurrence and distant recurrence states

The results of the DM sensitivity analysis using health state utilities derived from reported VAS utilities from de Almeida *et al.*[84] are reported in *Table 75*. Using the alternative utility values (for the recurrence states) has little impact on the results. The expected QALY values in both arms are slightly reduced compared with the base case, but the incremental effect is unchanged and PET–CT surveillance remains cost-effective.

Decision model sensitivity analysis 3: additional recurrences allowed beyond 5 years post diagnosis in local and distant recurrence health states

The results of DM sensitivity analysis 3 are presented in *Table 76*. Allowing primary and subsequent recurrences to occur beyond 5 years has little impact on the results. The QALY increment associated with PET–CT surveillance is slightly reduced (from 0.13 QALYs in the base case to 0.10 QALYs), which results in a small decrease in the expected INB (from 0.21 QALYs in the base case to 0.18 QALYs); however, PET–CT surveillance remains cost-effective. As in the base-case analysis, there is uncertainty around this result, with wide CIs around the expected health gain (0.18 QALYs, 95% CI –0.43 to 0.87 QALYs).

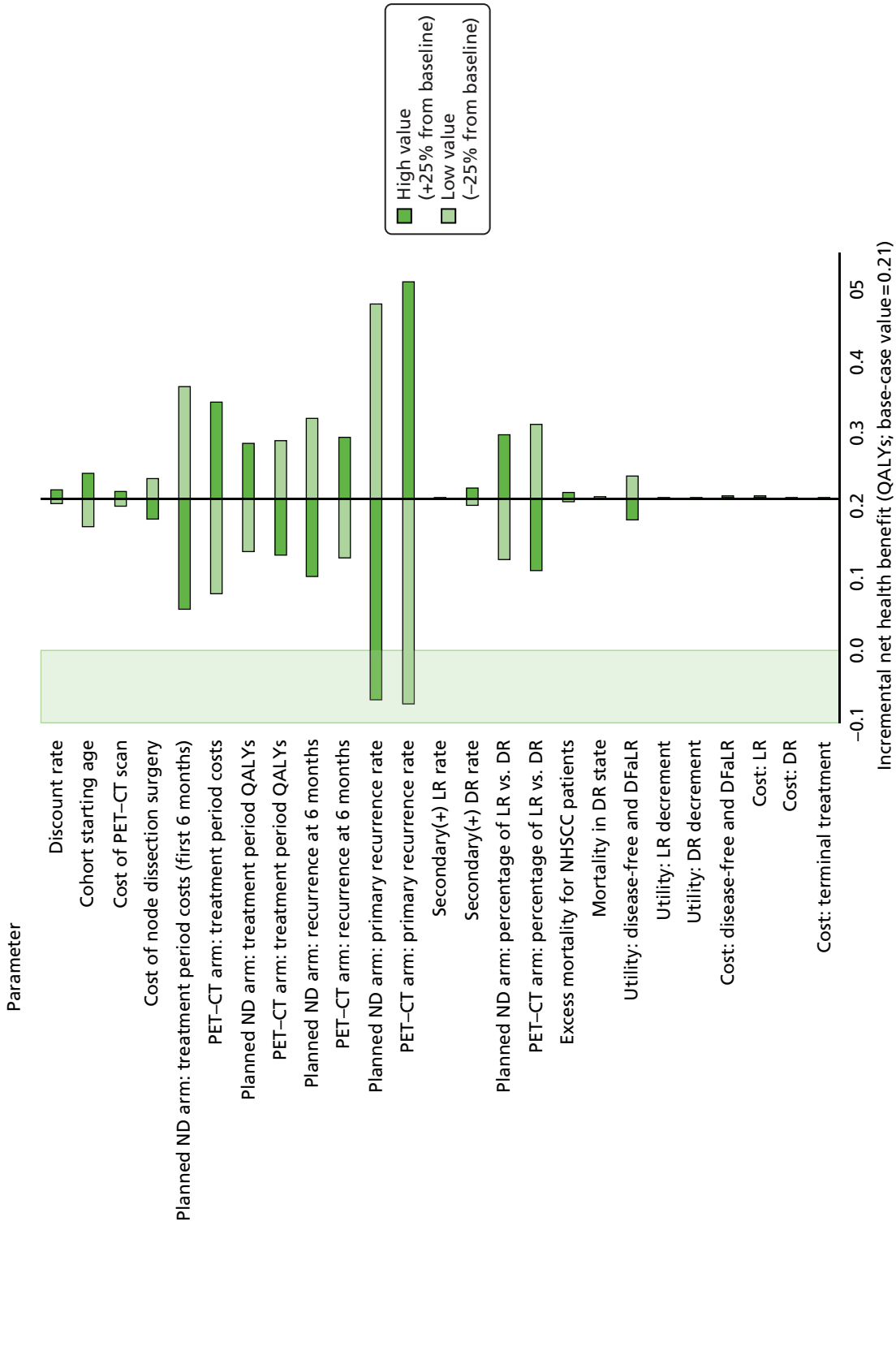


FIGURE 34 The results of DM one-way sensitivity analysis (tornado plot). DFaLR, disease free after local recurrence.

TABLE 73 Decision model sensitivity analysis 1.1: DM cost-effectiveness sensitivity analysis using additional cost data (NHS and PSS perspective, additional data used as reported)

Strategy	Total cost (£) (95% CI)	Total QALY (95% CI)	Incremental cost (£) (95% CI)	Incremental QALY (95% CI)	ICER	NB (QALYs) (95% CI)	Incremental NB (QALYs) (95% CI)
Planned ND	125,147 (91,267 to 166,177)	9.01 (7.88 to 10.47)	–	–	–	2.74 (0.50 to 4.65)	–
PET–CT	120,872 (86,905 to 161,955)	9.13 (8.05 to 10.54)	–4276 (–11,277 to 3154)	0.13 (–0.49 to 0.78)	Dominant	3.09 (0.88 to 5.04)	0.34 (0.02 to 0.70)

TABLE 74 Decision model sensitivity analysis 1.2: DM cost-effectiveness sensitivity analysis using additional cost data (NHS and PSS perspective, excluding secondary care resource use items identifying UHB or UHCW as the visited hospital)

Strategy	Total cost (£) (95% CI)	Total QALY (95% CI)	Incremental cost (£) (95% CI)	Incremental QALY (95% CI)	ICER	NB (QALYs) (95% CI)	Incremental NB (QALYs) (95% CI)
Planned ND	99,898 (68,360 to 139,654)	9.01 (7.87 to 10.46)	–	–	–	4.01 (1.90 to 5.87)	–
PET–CT	99,198 (67,304 to 139,049)	9.13 (8.05 to 10.54)	–700 (–6190 to 5362)	0.13 (–0.49 to 0.79)	Dominant	4.18 (2.04 to 6.04)	0.17 (–0.22 to 0.57)

TABLE 75 Decision model sensitivity analysis 2: DM cost-effectiveness sensitivity analysis – QALYs based on VAS utilities

Strategy	Total cost (£) (95% CI)	Total QALY (95% CI)	Incremental cost (£) (95% CI)	Incremental QALY (95% CI)	ICER	NB (QALYs) (95% CI)	Incremental NB (QALYs) (95% CI)
Planned ND	24,136 (12,977 to 63,956)	8.94 (7.85 to 10.09)	–	–	–	7.73 (5.60 to 9.10)	–
PET–CT	22,651 (11,323 to 62,958)	9.07 (8.02 to 10.09)	–1485 (–2805 to 223)	0.13 (–0.50 to 0.79)	Dominant	7.94 (5.78 to 9.28)	0.21 (–0.41 to 0.86)

TABLE 76 Decision model sensitivity analysis 3: DM cost-effectiveness sensitivity analysis – recurrences allowed beyond 5 years

Strategy	Total cost (£) (95% CI)	Total QALY (95% CI)	Incremental cost (£) (95% CI)	Incremental QALY (95% CI)	ICER	NB (QALYs) (95% CI)	Incremental NB (QALYs) (95% CI)
Planned ND	24,313 (13,149 to 63,642)	8.76 (7.60 – 10.32)	–	–	–	7.55 (5.41 to 9.12)	–
PET–CT	22,841 (11,492 to 62,359)	8.86 (7.73 – 10.40)	–1472 (–2934 to 381)	0.10 (–0.56 to 0.80)	Dominant	7.72 (5.55 to 9.29)	0.18 (–0.43 to 0.87)

Summary

An economic evaluation was conducted to assess the cost-effectiveness of PET–CT-guided watch-and-wait policy compared with planned ND. The evaluation consisted of two components: (1) a WT analysis, in which cost-effectiveness is assessed over the 24-month trial period using individual patient data collected in the trial; and (2) a decision-analytic model analysis, in which cost-effectiveness is assessed over a lifetime horizon, using standard modelling techniques applied to the trial data in order to extrapolate the trial results. Owing to the availability of data within the trial, the base-case analysis was conducted from a NHS secondary care perspective. Additional sensitivity analyses were conducted using a NHS and PSS perspective and a societal perspective (including patient and carer one-off expenses, travel costs and lost earnings).

Results of the WT base-case analysis indicate that the PET–CT watch-and-wait strategy is cost-effective over a 2-year time horizon. Using a NHS secondary care perspective, PET–CT surveillance yielded an average per-patient cost saving of £1492 (95% CI –£698 to –£2239) and an expected health gain of 0.08 QALYs (95% CI –0.03 to 0.19 QALYs) compared with planned ND. The PET–CT watch-and-wait policy therefore dominates standard care, being more effective and less costly, producing a 2-year incremental net health gain of 0.16 QALYs (95% CI 0.03 to 0.28 QALYs) per patient, with a 99% probability of being cost-effective at a £20,000 per QALY threshold.

The results of the base-case lifetime DM analysis indicate that PET–CT surveillance is expected to remain cost-effective over a lifetime horizon; however, there is increased uncertainty around this result compared with the WT analysis. In the base-case analysis the PET–CT watch-and-wait policy is associated with an expected lifetime cost saving of –£1485 (95% CI –£2815 to £159) compared with planned ND, and an expected health gain of 0.13 QALYs (95% CI –0.41 to 0.85 QALYs), resulting in an incremental net health gain of 0.21 QALYs (95% CI –0.41 to 0.85 QALYs). At a £20,000 per QALY threshold there is a 75% probability that PET–CT surveillance is the most cost-effective strategy, dropping to 68% at a £100,000 per QALY threshold. These results were sensitive to parameters relating to the rate of recurrence in each arm, although the incremental health benefit associated with PET–CT surveillance dropped below zero (i.e. indicating non-cost-effectiveness) in only two scenarios: (1) increasing the expected rate of primary recurrence in the PET–CT arm by 25% or (2) decreasing the rate of primary recurrence in the planned ND arm by 25%.

The results of both the WT and DM sensitivity analyses indicate that the inclusion of additional primary, secondary, community and societal care costs is likely to result in substantial increases in the expected cost of both treatment strategies, and may lead to an increase in the expected cost-effectiveness of PET–CT surveillance. However, these results should be considered with caution given the significant uncertainty involved in including the additional costs that were derived from small sample sizes from the trial.

Discussion

In the WT base-case analysis PET–CT surveillance was found to be cost-effective, producing an incremental cost saving of –£1415 and an incremental health benefit of 0.07 QALYs compared with planned ND over a 2-year time horizon. This result is relatively certain, with an expected INB of 0.15 QALYs (95% CI 0.02 to 0.27 QALYs) and a probability of cost-effectiveness of 99% at a £20,000 per QALY threshold, and with results being robust to sensitivity analyses. In the lifetime DM base-case analysis PET–CT surveillance remained cost-effective, producing a lifetime cost saving of –£1485 and an incremental health benefit of 0.13 QALYs. However, this result was more uncertain, with a wide CI around the expected incremental net benefit (0.21 QALYs, 95% CI –0.41 to 0.85 QALYs), and a probability of cost-effectiveness of 75% at a £20,000 per QALY threshold, dropping to 68% at a £100,000 per QALY threshold. This uncertainty is largely attributable to uncertainty around the QALY benefit associated with PET–CT surveillance (mean 0.13 QALYs, 95% CI –0.49 to 0.79 QALYs), which is likely to be a result of uncertainty around the relative rate of recurrences between the two arms. Although there is uncertainty around the lifetime results,

PET–CT surveillance remained cost-effective across a range of sensitivity analyses, and was found to be non-cost-effective only when assuming a significant increase in the rate of primary recurrences in the PET–CT arm compared with the planned ND arm, which is expected to represent a highly unlikely scenario. Thus, although uncertainty exists, the mean result was relatively robust to sensitivity analyses.

The results of this analysis are in line with previous economic evaluations of PET–CT surveillance strategies to guide the use of surgery in H&N cancer patients.[90–94] In a recent study, Pryor *et al.*[94] found that a similar PET–CT-guided strategy was a safe and significantly less costly alternative strategy than planned surgery from an Australian health service perspective.[94] Similarly, Rabalais *et al.*,[92] Sher *et al.*[90] and Hollenbeak *et al.*[93] all found PET–CT surveillance to be cost-effective from an American health-care perspective. As far as these authors are aware, no full economic evaluation of PET–CT surveillance for H&N cancer has previously been conducted from a UK NHS perspective; therefore, these results represent the first indication that PET–CT is likely to provide a cost-effective alternative to planned ND within the UK health-care system, and adds support to the current body of studies in favour of adopting PET–CT into routine clinical practice.

Ideally, cost-effectiveness analyses should be conducted from a NHS and PSS or societal perspective, in order to fully account for resources that will be consumed as a result of implementing the new intervention.[63] In this evaluation, however, we have adopted a NHS secondary care perspective in the base case, which accounts only for the use of secondary care (i.e. hospital) resources. This is a clear limitation of the analysis, and is a result of a lack of reliable and sufficient data within the trial on which to derive full NHS and PSS or societal costs; such data were available for only a small subsample of 42 patients and 35 carers. Sensitivity analyses were conducted to assess the potential impact of including these additional costs, the results of which indicate that inclusion of broader costs within the analysis is likely to result in substantial increases in the expected cost of both treatment strategies, and may in fact lead to PET–CT surveillance becoming more cost-effective than in the base case. This suggests that the use of PET–CT may lead to resource savings across primary and societal care, in addition to the savings evident in the hospital secondary care setting. As previously stated, these results should be interpreted with caution because of the uncertainty around the additional cost data used in these analyses; nevertheless, it is encouraging that these exploratory results indicate that PET–CT surveillance remains cost-effective when adopting a broader perspective. Future work should focus on obtaining more accurate estimates of NHS, PSS and societal care costs, in order to enable future UK economic evaluations to adopt a broader perspective.

Chapter 5 Discussion

Interpretation

The results of this study show that PET–CT-guided surveillance followed by ND appears to be equivalent to a planned ND regimen for the management of advanced (N2/N3) nodal disease in patients with H&N cancer being treated with CRT, and demonstrated non-inferiority for differences of > 4% in OS. This strongly supports the efficacy of the PET–CT-guided surveillance policy. The policy is also cost-effective in the short term, and potentially in the long term.

Surveillance policy resulted in far fewer patients requiring ND and, as a result, considerably fewer complications of surgery. The effect of ND on quality of life, however, was not high in that both groups appear to have similar quality of life in the longer term. It is notable, however, that patients randomised to the planned ND experienced more SAEs. One would expect these to be related to surgery, but these were mainly related to CRT, especially in the arm of patients who were randomised to ND after CRT. It is not clear why this should be, or whether or not this is a significant finding. The locoregional control rate and death from other causes were the same in both arms. Importantly, we demonstrated the feasibility and success of a PET–CT-guided policy in a multicentre randomised prospective setting, something that had not been demonstrated before.

The fact that concordance between the central laboratory and the local radiology readings was very high demonstrates the applicability of the surveillance regimen across PET–CT sites in the UK. It is also a testament to the quality of the training of PET–CT radiologists in the UK.

It is important to note that most of the patients had OPC and a large proportion of these were caused by the human papillomavirus (HPV-positive OPC). The subgroup analysis suggests that the surveillance policy is particularly effective in the HPV-positive subgroup. This may be because HPV-positive OPC responds well to CRT and, therefore, there is a smaller risk of persistent disease in HPV-negative patients. However, the subgroup analysis also shows the surveillance policy to be non-inferior for HPV-negative patients.

Generalisability

It is important to note that most of the patients had N2a or N2b disease and that very few patients had N3 disease. This may be partly because N3 disease is not common. However, there may also be a degree of selection bias by investigators, who may have been less keen to submit patients with N3 disease to randomisation. On the one hand, this may have been because they were concerned that the surveillance policy may result in residual disease in the N3 population, which would be difficult to salvage. On the other hand, clinicians may have been concerned that subjecting a patient to a planned ND, especially before CRT, may not be in the patient's best interest if there was a high chance that the primary site would not respond to CRT.

Overall evidence

The study supports claims that ND does not confer any survival benefits or an increase in nodal recurrence in patients who show a CR compared with patients who undergo a surveillance policy.[17,18]

The rate of complications following ND is only slightly higher following PET–CT than planned NDs done before or immediately after CRT. The differences in the rate of complications were not significant.

As was expected, outcomes were significantly better among patients with HPV-positive OPC than among those with HPV-negative disease. However, there is no difference between the planned ND and surveillance arms by HPV status. Indeed, HPV-positive patients may benefit from undergoing surveillance.

Of interest is the fact that there were no quality-of-life differences between the ND and surveillance arms. This suggests that the main determinant of quality of life in these patients is the CRT.

Also of interest is the fact that HPV-positive patients had significantly worse quality of life during treatment, but that they recovered from that and had an overall better quality of life than HPV-negative patients in the longer term. The reasons for this may be that the HPV-positive patients are a younger group who are usually employed and have young families; therefore, the effect of treatment may have more of an impact on their daily lives during that period. However, because they are younger and fitter, in the longer term they recover more quickly and have a higher baseline quality of life when considering the effect of age on quality of life.

Overall conclusions

In the authors' opinion, this study strongly supports the efficacy of the PET–CT-guided surveillance policy. The policy is also cost-effective in the short term, and potentially in the long term. Importantly, surveillance resulted in 80% of patients avoiding ND and, therefore, fewer complications of surgery. Patients in the surveillance arm had similar quality-of-life scores as patients in the ND arm and, when accounting for the cost-savings associated with PET–CT surveillance, the intervention was found to be more cost-effective in the short term and, potentially, in the long term also.

This is the largest report in the literature on the accuracy of PET–CT scanning in assessing residual disease at the primary site that provides sufficient evidence to allow the replacement of EUA (the current gold standard) by PET–CT scanning, which is a less invasive non-operative technique.

Most of the patients in this study had N2a or N2b disease and few patients had N3 disease; therefore, we cannot confirm the efficacy of the surveillance policy for very large N3 nodal disease.

Further research

The aim of the economic evaluation was to identify the WT and long-term incremental cost-effectiveness of a PET–CT-guided watch-and-wait policy compared with planned ND in HNSCC patients. As yet, comparisons of CT-driven response systems and PET–CT-driven systems in multicentre randomised settings have not been published. Unfortunately, because of cost constraints, we did not undertake contrast CT at the same time as PET–CT, so we cannot compare the efficacies of the two modalities. Therefore, the relative efficacy and cost-effectiveness of PET–CT compared with CT-based approaches is still unclear.

The majority of patients recruited into the study had N2 disease. This means that the study results and the PET–CT surveillance policy may not be applicable to very large N3 nodal disease. This requires further assessment.

Further research is also required for the equivocal cases. This is especially the case with PET-negative disease with persistent nodes on anatomical CT scanning. Research to assess prolonged surveillance with CT or assessment of the persistent node using ultrasound-guided fine-needle aspiration is warranted.

Acknowledgements

Trial Management Group

The Trial Management Group comprised Professor Hisham Mehanna (as chief investigator), Ms Joy Rahman (as trial co-ordinator), Mr Chris McConkey (as statistician), Professor Janet A Dunn (as senior statistician), Mr Dharmesh Patel (as trial administrator) and Ms Jennifer Cooper (as data input clerk).

Research nurses

June Jones and Lynda Wagstaff, research nurses, were funded by the trial to recruit patients at UHCW and Queen Elizabeth Hospital, Birmingham, respectively.

The Administration of Radioactive Substances Advisory Committee licence holders

The ARSAC licence holders for the study covered various PET–CT centres and included Dr Margaret Brookes for Aberdeen, Dr Anil Kumar for Basildon, Dr Julian Kabala for Bristol, Dr Peter Guest for Birmingham, Dr Jonathan Hill for Preston, Dr Richard Smith for Bradford, Dr Iain Lyburn for Linton House Clinic Cheltenham, Dr Paul Hulse for Christie, Dr Peter Guest for University Hospital Coventry and Warwickshire, Dr Rakesh Fanatra for Nottingham, Dr Dilip Patel for Edinburgh, Dr Thomas Gruning for Derriford, Dr Eddie Kalkman for West of Scotland, Dr Lucy Richardson for Hull, Dr Andrew Scarsbrook for Leeds, Dr Nuribe Wieshmann for Clatterbridge, Dr Michael O'Doherty and Dr Sally Barrington for St Thomas', Dr Gary Cook for Royal Marsden, Dr Irfan Kayani for University College London, Dr Wai-Lup Wong for Alliance Imaging London, Dr John Wilsdon for Newcastle, Dr Nicola Robson for Poole and Dr John Hall for Guildford Diagnostic Imaging Surrey.

Patient and public involvement

Mrs Ethel Culling and Ms Vivienne Reed of the Laryngectomy Society were members of the Trial Steering Committee and commented on the drafting of the protocol and the patient information sheets for the study.

The local head and neck teams

The PET-NECK team would like to thank the NHS trusts and participating principal investigators and their colleagues for recruiting and monitoring trial participants. These include (in alphabetical order of participating hospitals followed by the local principal investigators and their colleagues): Dr David Hurman of Aberdeen Royal Infirmary; Professor Terry Jones, Dr Hulya Weishmann, Mr Jeffrey Lancaster, Dr Aditya Shenoy, Ms Shirley Pringle and Mr Paul Banks of Aintree University Hospital; Mr Imtiaz Ahmed, Dr Sree Palvai, Dr Krinaswami Madhavan and Mrs Anne Hill of Basildon Hospital and Southend University Hospital; Dr Anna Thompson of Barnet and Chase Farm Hospitals; Mr Mohammed Rizwanullah of the Beatson West of Scotland Cancer Centre; Mr William Primrose of Belfast City Hospital and Royal Victoria Hospital; Dr Muthiah Sivaramalingham, Dr Ajay Mehta, Dr Ajay Nigam, Mr N Kazmi, Leanne Davies, Sue Hesketh, Alison Mackle, Karen Gratrix and Christopher Pemberton of Blackpool Victoria Hospital; Dr James McCaul, Andrew Pick, Shelley Berry, Raghav Kulkarni and Kayleigh Gilbert of Bradford Royal Infirmary; Dr Hoda Booz and Dr Matthew Beasley of the Bristol Haematology and Oncology Centre;

Dr Vivienne Loo of Broomfield Hospital; Dr Ashoke Biswas of Burnley General Hospital; Mr Jemy Jose of Castle Hill Hospital; Dr Audrey Cook of Cheltenham General Hospital and Gloucester Royal Hospital; Dr Om Purohit of Chesterfield Royal Hospital; Professor Jarrod Homer, Professor Nicholas Slevin, Dr Lip Wai Lee and Mrs Kim Denton of Christie Hospital; Dr Amy Roy of Derriford Hospital; Dr Elizabeth Junor and Mrs Kirsty Peebles of Dumfries and Galloway and Western General Hospital; Mr Jean-Pierre Jeannon and Dr Teresa Guerrerourbano of Guy's and St Thomas' Hospitals; Mr Richard Wight, Ms Rita Mohan, Mr Shane Lester, Dr Peter Dunlop, Mr Colin Edge, Mr Douglas Byant and Dr Eleanor Aynsley of James Cook University Hospital; Mr Alistair Balfour of Queen Elizabeth Queen Mother Hospital, Kent and Canterbury Hospital, and William Harvey Hospital; Dr Kate Goodchild and Ann Downs of Luton and Dunstable Hospitals, Mount Vernon Hospital and Northwick Park; Dr Caroline Brammer of New Cross Hospital; Dr Lip Lee, Joanne Allsop and Tracey Hodgkiss of North Manchester General Hospital; Dr Anna Thompson of North Middlesex Hospital; Dr Judy Christian and Sarah Fleet of Nottingham University Hospital; Dr Joseph Davies of Poole Hospital; Mr Kim Dae, Debi Barnes, Mr Costa Repanos, Dr Katy Bradley, Mr Tim Mellor, Mr Jean Daniel Dubois and Professor Peter Brennan of Queen Alexandra Hospital; Mr Keith Webster, Dr Andrew Hartley, Mrs Lynda Wagstaff, Dr Paul Sanghera, Dr John Glaholm, Mr Huw Griffiths, Mr Ijaz Ahmed, Mr Paul Pracey, Mr Christopher Jennings, Mr Sat Parmar, Mr Timothy Martin, Mrs Janet O'Connell, Dr Adrian Warfield, Dr Peter Colloby, Dr John Henderson, Dr Steve Colley, Dr Alison Page, Dr Julie Olliff, Dr Tunde Ogunremi and Dr Peter Guest of the Queen Elizabeth Hospital Birmingham; Dr Ashoke Biswas and Mr Pardeep Morar of Royal Blackburn Hospital; Mr Jerry Sharp of the Royal Derby Hospital; Dr David Hwang of the Royal Devon and Exeter Hospital; Professor Christopher Nutting of the Royal Marsden Hospital; Dr Muthiah Sivaramalingham, Dr Ajay Mehta, Dr Ashoke Biswas and Mr A Cardozo of the Royal Preston Hospital; Dr Stephen Whittaker, Dr Katie Wood and Mrs Ceri Jamieson of the Royal Surrey County Hospital; Dr Emma De Winton and Mr Tomas Tylee of the Royal United Hospital; Dr Prakash Ramachandra of Russells Hall Hospital; Dr Martin Rolles, Tracey Ford and Karen Phillips of the Singleton Hospital; Mr Vinidh Paleri and Dr Charles Kelly of the Sir Bobby Robson Cancer Trials Centre (Freeman Hospital); Mr Paul Tierney of Southmead Hospital; Mr James Moor, Eleanor Dungca, Ms Helen Cocks and Mr James O'Hara of Sunderland Royal Hospital; Dr Geoffrey Cogill of Torbay Hospital; Dr Martin Forster and Dr Dawn Carnell of University College Hospital; Professor Hisham Mehanna, Ms June Jones, Ms Rachel O'Beney, Mr Mark Wake, Mr Gary Walton, Mr Raj Sandhu, Dr Lydia Fresco, Dr Andrew Chan and Dr David Jones of UHCW; Dr Mererid Evans, Ms Amanda Jackson, Ms Caroline Vitolo, Mr Nachi Palaniappan, Ms Laura Moss, Mr Carl Badger Watts and Ms Catherine John of the Velindre Cancer Centre; Dr Andrew Hartley and Lynda Wagstaff of Walsall Manor Hospital; and Professor Martin Robinson, Dr Bernadette Foran, Gail Jessop and John Martindale of Weston Park Hospital.

The PET-NECK trial team would like to thank Jennifer Cooper (research team) for her contributions to the administration of the trial; Dr Bal Sanghera (Paul Strickland Scanner Centre) for performing quality checks on all the PET–CT scanners used in the trial; Dr Hulya Wieshmann for conducting a review of 10% of CT and MRI scans for QA; Dr Sandra Ventorin von Zeildler for undertaking the histopathology review for 10% of patients; the PET-NECK Collect team (Sean James, Andrew White and Adrian Fisk) for managing the translational aspects of the study; and Dr Max Robinson for carrying out p16 and HPV testing for trial participants. The team would also like to thank the Independent Data Monitoring Committee members (Dr Roger A'Hern, Professor Fons Balm and Professor Vincent Gregoire) and the Trial Steering Committee members (Professor Jonathan Ledermann, Professor Michael Langman, Dr Andrew Hartley, Professor Christopher Nutting, Professor Patrick Bradley, Ethel Culling and Vivien Reed) for their invaluable input and advice. Finally, we would like to thank the UK PET–CT centres for scanning patients and co-operating with the QA review, the providers of PET–CT scanning units (Inhealth, Alliance and Cobalt Healthcare) and the ARSAC licence holders for overseeing the safety aspects of patient PET–CT scanning.

Other acknowledgements

Alison F Smith, Peter Hall and Claire Hulme are supported by the National Institute for Health Research Diagnostic Evidence Co-operative Leeds.

This project benefited from facilities funded through the Department of Health: PET–CT scans performed for half the patients in the trial were paid for by the NHS. The project also received support from the National Institute for Health Research Clinical Research Networks. The PET-NECK team thank both sources for their valuable contributions.

Contributions of authors

Hisham Mehanna (Chief Investigator) and **Janet A Dunn** (Co-Applicant and Senior Statistician) were responsible for identifying the research question and writing the study protocol.

Chris C McConkey (Trial Statistician) and **Janet A Dunn** (Senior Statistician) carried out the analysis and interpreted the results.

Joy K Rahman (Trial Co-ordinator) managed the administration of the trial.

Wai-Lup Wong (Co-Applicant and Principal Radiologist) reviewed all trial PET–CT scans and advised hospital clinicians on treatment response.

Alison F Smith, **Peter Hall** and **Claire Hulme** (Health Economics Team) conducted the analysis of economic models for the trial.

Chris Nutting (Professor of Clinical Oncology) provided expertise on CR treatment.

Andrew GJ Hartley (Consultant Oncologist) provided advice on eligibility and treatment regimens.

Dharmesh K Patel (Trial Administrator) assisted with the administration of the trial.

Sandra Ventorin von Zeidler (Professor of Pathology) performed the histopathology review for 10% of patients and contributed to the writing of the manuscript.

Max Robinson (Senior Lecturer in Oral Pathology) undertook the p16 staining.

Bal Sanghera (Clinical Scientist) performed quality checks on all the PET–CT scanners used in the trial and contributed to the writing of the manuscript.

Lydia Fresco (Clinical Oncologist) recruited patients and is the clinical lead at the sponsor site.

Hisham Mehanna, **Chris C McConkey**, **Joy K Rahman**, **Alison F Smith** and **Janet A Dunn** were responsible for drafting this report, although all authors provided comments on drafts and approved the final version.

Publications

Rossignol M. Indications de la tomographie par émission de positrons (TEP): mise à jour sommaire. *ETMIS* 2011;**7**:1–44.

Kaur G, Hutchison I, Mehanna H, Williamson P, Shaw R, Tudur Smith C. Barriers to recruitment in surgical trials in head and neck oncology: a survey of trial investigators. *BMJ* 2013;**3**:e002625.

Mehanna H, Wong W-L, McConkey CC, Rahman JK, Robinson M, Hartley AGJ, *et al.* PET-NECK. A multi-centre, randomised, phase III, controlled trial (RCT) comparing PETCT guided active surveillance with planned neck dissection (ND) for locally advanced (N2/N3) nodal metastases (LANM) in patients with head and neck squamous cell cancer (HNSCC) treated with primary radical chemoradiotherapy (CRT). *J Clin Oncol* 2015;**33**. 2015 ASCO Annual Meeting, 29 May–2 June, Chicago, IL, abstract no. 6009.

Smith AF, Hall P, Hulme C, Dunn J, Rahman JK, McConkey CC, Mehanna H. Is PET-CT-guided management for patients with locally advanced head and neck squamous cell cancer (HNSCC) cost-effective? Results from a UK non-inferiority phase III randomized trial. *J Clin Oncol* 2015;**33**. 2015 ASCO Annual Meeting, 29 May–2 June, Chicago, IL, abstract no. 6010.

Mehanna H, Wong W-L, McConkey CC, Rahman JK, Robinson M, Hartley AGJ, *et al.* PET-CT surveillance versus neck dissection in advanced head and neck cancer. *N Engl J Med* 2016;**374**:1444–54.

Data sharing statement

All requests for data sharing will adhere to the Warwick Clinical Trials Unit data sharing agreement policy. Warwick Clinical Trials Unit is supportive of data sharing and will endeavour to assist in requests for data sharing. All requests are dealt with on a case-by-case basis. Any request should be submitted on the Warwick Clinical Trials Unit data request form, which is then reviewed by the Trial Management Group and the sponsor.

References

1. British Association of Otorhinolaryngologists Head and Neck Surgeons. *Effective Head and Neck Cancer Management*. Third Consensus Document. London: The Royal College of Surgeons; 2002.
2. Cooper JS, Pajak TF, Forastiere AA, Jacobs J, Campbell BH, Saxman SB, *et al.* Postoperative concurrent radiotherapy and chemotherapy for high-risk squamous-cell carcinoma of the head and neck. *N Engl J Med* 2004;**350**:1937–44. http://dx.doi.org/10.1056/NEJMoa032646
3. Bernier J, Domenge C, Ozsahin M, Matuszewska K, Lefèbvre JL, Greiner RH, *et al.* Postoperative irradiation with or without concomitant chemotherapy for locally advanced head and neck cancer. *N Engl J Med* 2004;**350**:1945–52. http://dx.doi.org/10.1056/NEJMoa032641
4. Pignon JP, le Maître A, Maillard E, Bourhis J, MACH-NC Collaborative Group. Meta-analysis of chemotherapy in head and neck cancer (MACH-NC): an update on 93 randomised trials and 17,346 patients. *Radiother Oncol* 2009;**92**:4–14. http://dx.doi.org/10.1016/j.radonc.2009.04.014
5. Thariat J, Hamoir M, Garrel R, Cosmidis A, Dassonville O, Janot, *et al.* Management of the neck in the setting of definitive chemoradiation: is there a consensus? A GETTEC study. *Ann Surg Oncol* 2012;**19**:2311–19. http://dx.doi.org/10.1245/s10434-012-2275-9
6. Gupta T, Agarwal JP. Planned neck dissection following chemo-radiotherapy in advanced HNSCC. *Int Semin Surg Oncol* 2004;**1**:6. http://dx.doi.org/10.1186/1477-7800-1-6
7. Kutler DI, Patel SG, Shah JP. The role of neck dissection following definitive chemoradiation. *Oncology* 2004;**18**:993–8.
8. Garg M, Beitler JJ. Controversies in management of the neck in head and neck cancer. *Curr Treat Options Oncol* 2004;**5**:35–40. https://doi.org/10.1007/s11864-004-0004-8
9. Mendenhall WM, Villaret DB, Amdur RJ, Hinerman RW, Mancuso AA. Planned neck dissection after definitive radiotherapy for squamous cell carcinoma of the head and neck. *Head Neck* 2002;**24**:1012–18. http://dx.doi.org/10.1002/hed.10187
10. Carinci F, Cassano L, Farina A, Pelucchi S, Calearo C, Modugno V, *et al.* Unresectable primary tumor of head and neck: does neck dissection combined with chemoradiotherapy improve survival? *J Craniofac Surg* 2001;**12**:438–43. https://doi.org/10.1097/00001665-200109000-00007
11. Lavertu P, Adelstein DJ, Saxton JP, Secic M, Wanamaker JR, Eliachar I, *et al.* Management of the neck in a randomized trial comparing concurrent chemotherapy and radiotherapy with radiotherapy alone in resectable stage III and IV squamous cell head and neck cancer. *Head Neck* 1997;**109**:559–66. https://doi.org/10.1002/(SICI)1097-0347(199710)19:7<559::AID-HED1>3.0.CO;2-6
12. McHam SA, Adelstein DJ, Rybicki LA, Lavertu P, Esclamado RM, Wood BG, *et al.* Who merits a neck dissection after definitive chemoradiotherapy for N2-N3 squamous cell head and neck cancer? *Head Neck* 2003;**25**:791–8. http://dx.doi.org/10.1002/hed.10293
13. Brizel DM, Prosnitz RG, Hunter S, Fisher SR, Clough RL, Downey MA, Scher RL. Necessity for adjuvant neck dissection in setting of concurrent chemoradiation for advanced head-and-neck cancer. *Int J Radiat Oncol Biol Phys* 2004;**58**:1418–23. http://dx.doi.org/10.1016/j.ijrobp.2003.09.004
14. Frank DK, Hu KS, Culliney BE, Persky MS, Nussbaum M, Schantz SP, *et al.* Planned neck dissection after concomitant radiochemotherapy for advanced head and neck cancer. *Laryngoscope* 2005;**115**:1015–20. https://doi.org/10.1097/01.MLG.0000162648.37638.76

15. Stenson KM, Haraf DJ, Pelzer H, Recant W, Kies MS, Weichselbaum RR, Vokes EE. The role of cervical lymphadenectomy after aggressive concomitant chemoradiotherapy: the feasibility of selective neck dissection. *Arch Otolaryngol Head Neck Surg* 2000;**126**:950–6. https://doi.org/10.1001/archotol.126.8.950

16. Clayman GL, Johnson CJ, Morrison W, Ginsberg L, Lippman SM. The role of neck dissection after chemoradiotherapy for oropharyngeal cancer with advanced nodal disease. *Arch Otolaryngol Head Neck Surg* 2001;**127**:135–9. https://doi.org/10.1001/archotol.127.2.135

17. Grabenbauer GG, Rödel C, Ernst-Stecken A, Brunner T, Hornung J, Kittel K, *et al.* Neck dissection following radiochemotherapy of advanced head and neck cancer – for selected cases only? *Radiother Oncol* 2003;**66**:57–63. https://doi.org/10.1016/S0167-8140(02)00193-7

18. Argiris A, Stenson KM, Brockstein BE, Mittal BB, Pelzer H, Kies MS, *et al.* Neck dissection in the combined-modality therapy of patients with locoregionally advanced head and neck cancer. *Head Neck* 2004;**26**:447–55. http://dx.doi.org/10.1002/hed.10394

19. Thariat J, Ramus L, Maingon P, Odin G, Gregoire V, Darcourt V, *et al.* Dental maps: automatic dental delineation for radiotherapy planning in head-and-neck cancer. *Int J Radiat Oncol Biol Phys* 2012;**82**:1858–65. http://dx.doi.org/10.1016/j.ijrobp.2011.03.035

20. Strasser MD, Gleich LL, Miller MA, Saavedra HI, Gluckman JL. Management implications of evaluating the N2 and N3 neck after organ preservation therapy. *Laryngoscope* 1999;**109**:1776–80. http://dx.doi.org/10.1097/00005537-199911000-00010

21. Isles MG, McConkey C, Mehanna HM. A systematic review and meta-analysis of the role of positron emission tomography in the follow up of head and neck squamous cell carcinoma following radiotherapy or chemoradiotherapy. *Clin Otolaryngol* 2008;**33**:210–22. https://doi.org/10.1111/j.1749-4486.2008.01688.x

22. Gupta T, Master Z, Kannan S, Agarwal JP, Ghsoh-Laskar S, Rangarajan V, *et al.* Diagnostic performance of post-treatment FDG PET or FDG PET/CT imaging in head and neck cancer: a systematic review and meta-analysis. *Eur J Nucl Med Mol Imaging* 2011;**38**:2083–95. http://dx.doi.org/10.1007/s00259-011-1893-y

23. Wong WL, Chevretton EB, McGurk M, Hussain K, Davis J, Beaney R, *et al.* A prospective study of PET-FDG imaging for the assessment of head and neck squamous cell carcinoma. *Clin Otolaryngol Allied Sci* 1997;**22**:209–14. https://doi.org/10.1046/j.1365-2273.1997.00852.x

24. Wong RJ, Lin DT, Schöder H, Patel SG, Gonen M, Wolden S, *et al.* Diagnostic and prognostic value of [(18)F]fluorodeoxyglucose positron emission tomography for recurrent head and neck squamous cell carcinoma. *J Clin Oncol* 2002;**20**:4199–208. https://doi.org/10.1200/JCO.2002.02.590

25. Fischbein NJ, AAssar OS, Caputo GR, Kaplan MJ, Singer MI, Price DC, *et al.* Clinical utility of positron emission tomography with 18F-fluorodeoxyglucose in detecting residual/recurrent squamous cell carcinoma of the head and neck. *AJNR Am J Neuroradiol* 1998;**19**:1189–96.

26. Greven KM. Positron-emission tomography for head and neck cancer. *Semin Radiat Oncol* 2004;**14**:121–9. https://doi.org/10.1053/j.semradonc.2003.12.005

27. Kubota K, Yokoyama J, Yamaguchi K, Ono S, Qureshy A, Itoh M, *et al.* FDG-PET delayed imaging for the detection of head and neck cancer recurrence after radio-chemotherapy: comparison with MRI/CT. *Eur J Nucl Med Mol Imaging* 2004;**3**:590–5. https://doi.org/10.1007/s00259-003-1408-6

28. Schmidt M, Schmalenbach M, Jungehülsing M, Theissen P, Dietlein M, Schröder U, *et al.* 18F-FDG PET for detecting recurrent head and neck cancer, local lymph node involvement and distant metastases. Comparison of qualitative visual and semiquantitative analysis. *Nuklearmedizin* 2004;**43**:91–101. http://dx.doi.org/10.1267/NUKL04030091

29. Yao M, Graham MM, Hoffman HT, Smith RB, Funk GF, Graham SM, *et al.* The role of post-radiation therapy FDG PET in prediction of necessity for post-radiation therapy neck dissection in locally advanced head-and-neck squamous cell carcinoma. *Int J Radiat Oncol Biol Phys* 2004;**59**:1001–10. http://dx.doi.org/10.1016/j.ijrobp.2004.01.040

30. Ryan WR, Fee WE, Le QT, Pinto HA. Positron-emission tomography for surveillance of head and neck cancer. *Laryngoscope* 2005;**115**:645–50. https://doi.org/10.1097/01.mlg.0000161345.23128.d4

31. Goerres GW, Schmid DT, Bandhauer F, Huguenin PU, von Schulthess GK, Schmid S, Stoeckli SJ. Positron emission tomography in the early follow-up of advanced head and neck cancer. *Arch Otolaryngol Head Neck Surg* 2004;**130**:105–9. http://dx.doi.org/10.1001/archotol.130.1.105

32. Hanasono MM, Kunda LD, Segall GM, Ku GH, Terris DJ. Uses and limitations of FDG positron emission tomography in patients with head and neck cancer. *Laryngoscope* 1999;**109**:880–5. https://doi.org/10.1097/00005537-199906000-00007

33. Branstetter BF, Blodgett TM, Zimmer LA, Snyderman CH, Johnson JT, Raman S, Meltzer CC. Head and neck malignancy: is PET/CT more accurate than PET or CT alone? *Radiology* 2005;**235**:580–6. http://dx.doi.org/10.1148/radiol.2352040134

34. Rusthoven KE, Koshy M, Paulino AC. The role of PET-CT fusion in head and neck cancer. *Oncology* 2005;**19**:241–6.

35. Terhaard CH, Bongers V, van Rijk PP, Hordijk GJ. F-18-fluoro-deoxy-glucose positron-emission tomography scanning in detection of local recurrence after radiotherapy for laryngeal/ pharyngeal cancer. *Head Neck* 2001;**23**:933–41. https://doi.org/10.1002/hed.1135

36. Lowe VJ, Boyd JH, Dunphy FR, Kim H, Dunleavy T, Collins BT, *et al.* Surveillance for recurrent head and neck cancer using positron emission tomography. *J Clin Oncol* 2000;**18**:651–8.

37. Bongers V, Hobbelink MG, van Rijk PP, Hordijk GJ. Cost-effectiveness of dual-head 18F-fluorodeoxyglucose PET for the detection of recurrent laryngeal cancer. *Cancer Biother Radiopharm* 2002;**17**:303–6. http://dx.doi.org/10.1089/10849780260179260

38. Brouwer J, Bodar EJ, De Bree R, Langendijk JA, Castelijns JA, Hoekstra OS, Leemans CR. Detecting recurrent laryngeal carcinoma after radiotherapy: room for improvement. *Eur Arch Otorhinolaryngol* 2004;**261**:417–22. http://dx.doi.org/10.1007/s00405-003-0708-6

39. Stokkel MP, Terhaard CH, Hordijk GJ, van Rijk PP. The detection of local recurrent head and neck cancer with fluorine-18 fluorodeoxyglucose dual-head positron emission tomography. *Eur J Nucl Med* 1999;**26**:767–73. https://doi.org/10.1007/s002590050448

40. Mehanna H, Olaleye O, Licitra L. Oropharyngeal cancer – is it time to change management according to human papilloma virus status? *Curr Opin Otolaryngol Head Neck Surg* 2012;**20**:120–4. http://dx.doi.org/10.1097/MOO.0b013e3283509735

41. Ang KK, Harris J, Wheeler R, Weber R, Rosenthal DI, Nguyen-Tân PF, *et al.* Human papillomavirus and survival of patients with oropharyngeal cancer. *N Engl J Med* 2010;**363**:24–35. http://dx.doi.org/10.1056/NEJMoa0912217

42. Forastiere AA, Goepfert H, Maor M, Pajak TF, Weber R, Morrison W, *et al.* Concurrent chemotherapy and radiotherapy for organ preservation in advanced laryngeal cancer. *N Engl J Med* 2003;**349**:2091–8. http://dx.doi.org/10.1056/NEJMoa031317

43. Great Britain. *Data Protection Act 2004*. London: The Stationery Office; 2004.

44. Calais G, Alfonsi M, Bardet E, Sire C, Germain T, Bergerot P, *et al.* Randomized trial of radiation therapy versus concomitant chemotherapy and radiation therapy for advanced-stage oropharynx carcinoma. *J Natl Cancer Inst* 1999;**91**:2081–6. https://doi.org/10.1093/jnci/91.24.2081

45. Cancer Action Team. *Head and Neck Measures for the Manual for Cancer Services*. 2006. URL: www.cquins.nhs.uk (accessed 26 January 2017).

46. Helliwell T, Woolgar J. *Minimum Dataset for Histopathology Reports on Head and Neck Carcinomas and Salivary Neoplasms*. London: The Royal College of Pathologists; 2005.

47. Great Britain. *Ionising Radiation (Medical Exposure) Regulations 2000*. London: The Stationery Office; 2000.

48. Sobin LH, Wittekind C, editors. *TNM Classification of Malignant Tumours*. 6th edn. New York, NY: John Wiley & Sons; 2002.

49. Wee J, Tan EH, Tai BC, Wong HB, Leong SS, Tan T, *et al.* Randomized trial of radiotherapy versus concurrent chemoradiotherapy followed by adjuvant chemotherapy in patients with American Joint Committee on Cancer/International Union against cancer stage III and IV nasopharyngeal cancer of the endemic variety. *J Clin Oncol* 2005;**23**:6730–8. http://dx.doi.org/10.1200/JCO.2005.16.790

50. Bachaud JM, Cohen-Jonathan E, Alzieu C, David JM, Serrano E, Daly-Schveitzer N. Combined postoperative radiotherapy and weekly cisplatin infusion for locally advanced head and neck carcinoma: final report of a randomized trial. *Int J Radiat Oncol Biol Phys* 1996;**36**:999–1004. https://doi.org/10.1016/S0360-3016(96)00430-0

51. Bonner JA, Harari PM, Giralt J, Azarnia N, Shin DM, Cohen RB, *et al.* Radiotherapy plus cetuximab for squamous-cell carcinoma of the head and neck. *N Engl J Med* 2006;**354**:567–78. http://dx.doi.org/10.1056/NEJMoa053422

52. Bernier J, Schneider D. Cetuximab combined with radiotherapy: an alternative to chemoradiotherapy for patients with locally advanced squamous cell carcinomas of the head and neck? *Eur J Cancer* 2007;**43**:35–45. http://dx.doi.org/10.1016/j.ejca.2006.08.035

53. Remenar E, Van Herpen C, Germa Lluch J, Stewart S, Gorlia T, Degardin M, *et al.* A randomized phase III multicenter trial of neoadjuvant docetaxel plus cisplatin and 5-fluorouracil (TPF) versus neoadjuvant PF in patients with locally advanced unresectable squamous cell carcinoma of the head and neck (SCCHN). Final analysis of EORTC 24971. *J Clin Oncol* 2006;**24**:5516.

54. Scottish Intercollegiate Guidelines Network. *Diagnosis and Management of Head and Neck Cancer: Guideline No. 90*. Edinburgh: Scottish Intercollegiate Guidelines Network; 2006.

55. Health and Safety Executive. *Health and Safety Guidance HSG95. Radiation Safety of Lasers used for Display Purposes*. Merseyside: Health and Safety Executive; 1996.

56. Moher D, Schulz K, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet* 2001;**357**:1191–4. https://doi.org/10.1016/S0140-6736(00)04337-3

57. Early Breast Cancer Trialists' Collaborative Group. *Treatment of Early Breast Cancer. Volume 1. Worldwide Evidence 1985–1990*. Oxford: Oxford University Press; 1990.

58. Gray RJ. A class of K-sample tests for comparing the cumulative incidence of a competing risk. *Ann Stat* 1988;**16**:1141–54. https://doi.org/10.1214/aos/1176350951

59. Bjordal K, de Graeff A, Fayers PM, Hammerlid E, van Pottelsberghe C, Curran D, *et al.* A 12 country field study of the EORTC QLQ-C30 (version 3.0) and the head and neck cancer specific module (EORTC QLQ-H&N35) in head and neck patients. EORTC Quality of Life Group. *Eur J Cancer* 2000;**36**:1796–807. https://doi.org/10.1016/S0959-8049(00)00186-6

60. Fayers PM. Interpreting quality of life data: population-based reference data for the EORTC QLQ-C30. *Eur J Cancer* 2001;**37**:1331–4. https://doi.org/10.1016/S0959-8049(01)00127-7

61. Chen AY, Frankowski R, Bishop-Leone J, Hebert T, Leyk S, Lewin J, Goepfert H. The Development and Validation of a dysphagia-specific quality-of-life questionnaire for patients with head and neck cancer. *Arch Otolaryngol Head Neck Surg* 2001;**127**:870–6.

62. Mehanna H, Wong W-L, McConkey CC, Rahman JK, Robinson M, Hartley AGJ, *et al.* PET-CT surveillance versus neck dissection in advanced head and neck cancer. *N Engl J Med* 2016;**374**:1444–54. http://dx.doi.org/10.1056/NEJMoa1514493

63. NICE. *Guide to the Methods of Technology Appraisal 2013*. NICE Process (PMG9). London: NICE; 2013.

64. EuroQol Research Foundation. *EuroQol*. Rotterdam: EuroQol Research Foundation; 2015. URL: www.euroqol.org (accessed 23 September 2016).

65. Dolan P. Modeling valuations for EuroQol health states. *Med Care* 1997;**35**:1095–108. https://doi.org/10.1097/00005650-199711000-00002

66. Wilcoxon F. Individual comparisons by ranking methods. *Biometrics Bulletin* 1945;**1**:80–3. https://doi.org/10.2307/3001968

67. Billingham LJ, Abrams KR. Simultaneous analysis of quality of life and survival data. *Stat Methods Med Res* 2002;**11**:25–48. https://doi.org/10.1191/0962280202sm269ra

68. Associate Director of Operations. *Performance Report Month 11*. March 2014 Trust Board Meeting Stevenage: East and North Hertfordshire NHS Trust; 2014.

69. Department of Health. *Reference Costs 2013–2014*. London: Department of Health; 2014.

70. Curtis L. *Unit Costs of Health and Social Care*. Canterbury: Personal Social Services Research Unit, University of Kent; 2014.

71. Department of Health. *Drugs and Pharmaceutical Electronic Market Information (eMit)*. London: Department of Health; 2015.

72. Joint Formulary Committee. *British National Formulary* (online). London: BMJ Group and Pharmaceutical Press. URL: www.medicinescomplete.com (accessed April 2015).

73. Hall P, Walkington L, Newsham A, Hall G, Glaser A, editors. Costs of hospital care over ten years from diagnosis of early breast cancer in England. *Eur J Cancer* 2014;**50**(Suppl. 2):S79–80.

74. Office for National Statistics. *Annual Survey of Hours and Earnings, 2014 Provisional Results*. London: Office for National Statistics; 2014.

75. NHS Employers. *Agenda for Change Pay*. 2014. URL: www.nhsemployers.org/your-workforce/pay-and-reward/pay/agenda-for-change-pay (accessed April 2015).

76. Office for National Statistics. *Rate of Inflation Joint Lowest Since Records Began*. London: Office for National Statistics; 2015.

77. Stinnett AA, Mullahy J. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. *Med Decis Making* 1998;**18**(Suppl. 2):68–80. https://doi.org/10.1177/0272989X98018002S09

78. Schliephake H, Jamil MU. Prospective evaluation of quality of life after oncologic surgery for oral cancer. *Int J Oral Maxillofac Surg* 2002;**31**:427–33. https://doi.org/10.1054/ijom.2001.0194

79. Hammerlid E, Silander E, Hörnestam L, Sullivan M. Health-related quality of life three years after diagnosis of head and neck cancer – a longitudinal study. *Head Neck* 2001;**23**:113–25. https://doi.org/10.1002/1097-0347(200102)23:2<113::AID-HED1006>3.0.CO;2-W

80. Chandu A, Sun KC, DeSilva RN, Smith AC. The assessment of quality of life in patients who have undergone surgery for oral cancer: a preliminary report. *J Oral Maxillofac Surg* 2005;**63**:1606–12. http://dx.doi.org/10.1016/j.joms.2005.07.012

81. De Graeff A, De Leeuw J, Ros W, Hordijk G, Blijham G, Winnubst J. A prospective study on quality of life of patients with cancer of the oral cavity or oropharynx treated with surgery with or without radiotherapy. *Oral Oncol* 1999;**35**:27–32. https://doi.org/10.1016/S1368-8375(98)00049-9

82. Parthan A, Posner MR, Brammer C, Beltran P, Jansen JP. Cost utility of docetaxel as induction chemotherapy followed by chemoradiation in locally advanced squamous cell carcinoma of the head and neck. *Head Neck* 2009;**31**:1255–62. http://dx.doi.org/10.1002/hed.21096

83. Brown B, Diamantopoulos A, Bernier J, Schöffski P, Hieke K, Mantovani L, *et al.* An economic evaluation of cetuximab combined with radiotherapy for patients with locally advanced head and neck cancer in Belgium, France, Italy, Switzerland, and the United Kingdom. *Value Health* 2008;**11**:791–9.

84. de Almeida JR, Villanueva NL, Moskowitz AJ, Miles BA, Teng MS, Sikora A, *et al.* Preferences and utilities for health states after treatment for oropharyngeal cancer: transoral robotic surgery versus definitive (chemo)radiotherapy. *Head Neck* 2014;**36**:923–33. http://dx.doi.org/10.1002/hed.23340

85. Office for National Statistics. *How Have Mortality Rates by Age Changed over the Last 50 Years?* London: Office for National Statistics. 2014. URL: www.ons.gov.uk/ons/rel/vsob1/death-reg-sum-tables/2013/sty-mortality-rates-by-age.html (accessed January 2015).

86. van der Schroeff MP, van de Schans SA, Piccirillo JF, Langeveld TP, Baatenburg de Jong RJ, Janssen-Heijnen ML. Conditional relative survival in head and neck squamous cell carcinoma: Permanent excess mortality risk for long-term survivors. *Head Neck* 2010;**32**:1613–18. http://dx.doi.org/10.1002/hed.21369

87. Wong LY, Wei WI, Lam LK, Yuen AP. Salvage of recurrent head and neck squamous cell carcinoma after primary curative surgery. *Head Neck* 2003;**25**:953–9. http://dx.doi.org/10.1002/hed.10310

88. Matoscevic K, Graf N, Pezier TF, Huber GF. Success of salvage treatment: a critical appraisal of salvage rates for different subsites of HNSCC. *Otolaryngol Head Neck Surg* 2014;**151**:454–61. http://dx.doi.org/10.1177/0194599814535183

89. Briggs A, Sculpher M, Claxton K. *Decision Modelling for Health Economic Evaluation*. Oxford: Oxford University Press; 2006.

90. Sher DJ, Tishler RB, Annino D, Punglia RS. Cost-effectiveness of CT and PET-CT for determining the need for adjuvant neck dissection in locally advanced head and neck cancer. *Ann Oncol* 2010;**21**:1072–7. http://dx.doi.org/10.1093/annonc/mdp405

91. van Hooren AC, Brouwer J, de Bree R, Hoekstra OS, Leemans CR, Uyl-de Groot CA. The cost-effectiveness of 18FDG-PET in selecting patients with suspicion of recurrent laryngeal carcinoma after radiotherapy for direct laryngoscopy. *Eur Arch Otorhinolaryngol* 2009;**266**:1441–8. http://dx.doi.org/10.1007/s00405-008-0878-3

92. Rabalais A, Walvekar RR, Johnson JT, Smith KJ. A cost-effectiveness analysis of positron emission tomography-computed tomography surveillance versus up-front neck dissection for management of the neck for N2 disease after chemoradiotherapy. *Laryngoscope* 2012;**122**:311–14. http://dx.doi.org/10.1002/lary.22464

93. Hollenbeak CS, Lowe VJ, Stack BC. The cost-effectiveness of fluorodeoxyglucose 18-F positron emission tomography in the N0 neck. *Cancer* 2001;**92**:2341–8. https://doi.org/10.1002/1097-0142(20011101)92:9<2341::AID-CNCR1581>3.0.CO;2-8

94. Pryor DI, Porceddu SV, Scuffham PA, Whitty JA, Thomas PA, Burmeister BH. Economic analysis of FDG-PET-guided management of the neck after primary chemoradiotherapy for node-positive head and neck squamous cell carcinoma. *Head Neck* 2013;**35**:1287–94. http://dx.doi.org/10.1002/hed.23108

EME
HS&DR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

Published by the NIHR Journals Library